



Universitat d'Alacant
Universidad de Alicante

ELABORACIÓN Y VALIDACIÓN DE UN
CUESTIONARIO SOBRE SINTOMATOLOGÍA
ASOCIADA A LAS DISFUNCIONES VISUALES

Mario Cantó Cerdán



Tesis **Doctorales**

UNIVERSIDAD de ALICANTE

Unitat de Digitalització UA
Unidad de Digitalización UA



Universitat d'Alacant
Universidad de Alicante

Departamento de Óptica, Farmacología y Anatomía

Facultad de Ciencias

**ELABORACIÓN Y VALIDACIÓN DE UN
CUESTIONARIO SOBRE SINTOMATOLOGÍA
ASOCIADA A LAS DISFUNCIONES VISUALES**

MARIO CANTÓ CERDÁN

Tesis presentada para aspirar al grado de
DOCTOR POR LA UNIVERSIDAD DE ALICANTE

Universitat d'Alacant
Universidad de Alicante
Doctorado en Ciencias de la Salud

Dirigida por:

PILAR CACHO MARTÍNEZ

Profesora Titular de Universidad

Departamento de Óptica, Farmacología y Anatomía

y

ÁNGEL GARCÍA MUÑOZ

Profesor Titular de Universidad

Departamento de Óptica, Farmacología y Anatomía

AGRADECIMIENTOS

En primer lugar, me gustaría agradecer a mi padre y a mi madre todo su esfuerzo y apoyo que han realizado toda la vida por toda mi formación académica, y, sobre todo, por toda su educación. No cabe en estas líneas todo lo que les debo, además de la vida. Cómo no, también quiero agradecer a mi hermana por su interés y momentos de apoyo. Gracias Marina.

Agradecer a todos mis amigos, Ángel, Dani, José, Nacho y Víctor, aunque siempre en la distancia debido a que las situaciones de la vida nos han separado, nunca habéis fallado cuando habéis tenido que estar ahí. Y sobre todo a Esther, siempre has estado ahí apoyándome, en lo bueno y en lo malo, ojalá sea así siempre.

A mis compañeras y compañeros de trabajo de Vissum Miranza Alicante, su interés y apoyo siempre ha sido un motivo para no rendirme en los momentos difíciles. Desde el momento que llegué hasta el día de hoy, solo puedo darles las gracias por todo lo que han hecho por mí tanto en lo profesional como en lo personal, gracias por vuestra confianza.

Agradecer a Paco Lara, y a todo el equipo de Óptica Lara, por su ayuda y colaboración en esta investigación, estaré siempre infinitamente agradecido y dispuesto a ayudar y colaborar en lo que se necesite. También agradecer a todas las personas que han participado en el estudio.

Por último, y más especial, me gustaría agradecer a Ángel y Pilar, directores de esta tesis doctoral, por toda su infinita paciencia, ayuda, y dedicación en este largo camino. Durante todos estos años en los que me he formado como optometrista,

siempre he tenido claro que sois mi modelo a seguir, y el espejo en donde mirarme. Aun me queda mucho para llegar a ser como vosotros, pero ojalá lo llegue a conseguir. Sé que hemos tenido muchas dificultades para llegar hasta el final en esta investigación, y sin vosotros no hubiese sido posible. Solo me quedo con lo bueno, y si tuviera que volver atrás y volver a elegir, os volvería a elegir a vosotros. Muchas gracias.

Embrace

Your

Dreams

Universitat d'Alacant
Universidad de Alicante



INDICE

SECCIÓN 1: SÍNTESIS DE LA TESIS	1
Resumen	3
1.1. Antecedentes y estado actual del tema	7
1.1.1. Introducción	7
1.1.2. Disfunciones visuales	14
1.1.3. Síntomas visuales	19
1.1.4. Herramientas para la recogida de síntomas visuales	22
1.1.5. Desarrollo y validación de cuestionarios	23
1.1.5.1. Identificación de ítems	24
1.1.5.2. Selección de los ítems	26
1.1.5.3. Análisis psicométrico	27
1.1.5.4. Exactitud diagnóstica mediante curvas ROC	36
1.1.5.5. Validez	36
1.1.5.6. Repetibilidad	38
1.2. Objetivos	39
1.3. Hipótesis	41
1.4. Justificación	43
1.5. Resumen global de los resultados obtenidos y discusión	45
1.5.1. Metodología Delphi para la detección de sintomatología asociada a las disfunciones visuales	47
1.5.1.1. Primera ronda de la metodología Delphi	49
1.5.1.2. Segunda ronda de la metodología Delphi	57
1.5.1.3. Tercera ronda de la metodología Delphi	60
1.5.2. Análisis Rasch para el desarrollo y reducción del Symptom Questionnaire for Visual Dysfunctions (SQVD)	67
1.5.2.1. Análisis de las curvas de probabilidad de respuesta	69

1.5.2.2. Reducción psicométrica de los ítems	71
1.5.2.3. Ajuste de los ítems al modelo usando los estadísticos infit y outfit	73
1.5.2.4. Análisis del funcionamiento diferencial de los ítems (DIF)	75
1.5.2.5. Fiabilidad de la separación de las medidas de las personas y de los ítems	76
1.5.2.6. Unidimensionalidad	76
1.5.2.7. Medida del alineamiento entre ítems y personas (targeting)	77
1.5.2.8. Tabla de transformación de las puntuaciones	79
1.5.3. Validación del Symptom Questionnaire for Visual Dysfunctions (SQVD): un cuestionario para evaluar los síntomas de cualquier tipo de disfunción visual	83
1.5.3.1. Análisis de las curvas de probabilidad de respuesta	86
1.5.3.2. Ajuste de los ítems al modelo usando los estadísticos infit y outfit	88
1.5.3.3. Análisis del funcionamiento diferencial de los ítems (DIF)	89
1.5.3.4. Fiabilidad de la separación de las medidas de las personas y de los ítems	90
1.5.3.5. Medida del alineamiento entre ítems y personas (targeting)	91
1.5.3.6. Unidimensionalidad y dependencia local	92
1.5.3.7. Tabla de transformación de las puntuaciones	94
1.5.3.8. Exactitud diagnóstica mediante curvas ROC	96
1.5.3.9. Validez	96
1.5.3.10. Repetibilidad	98

SECCIÓN 2: TRABAJOS PUBLICADOS	103
2.1. Capítulo I: Delphi methodology for symptomatology associated with visual dysfunctions	105
2.2. Capítulo II: Rasch analysis for development and reduction of symptom questionnaire for visual dysfunctions (SQVD)	121
2.3. Capítulo III: Validation of the symptom questionnaire for visual dysfunctions (SQVD): a questionnaire to evaluate symptoms of any type of visual dysfunctions	133
SECCIÓN 3: CONCLUSIONES	147
Conclusiones	149
SECCIÓN 4: LÍNEAS FUTURAS DE INVESTIGACIÓN	151
SECCIÓN 5: BIBLIOGRAFÍA	155
SECCIÓN 6: ANEXOS	175

SECCIÓN 1:

SÍNTESIS DE LA TESIS

Universitat d'Alacant
Universidad de Alicante

RESUMEN

La presencia de cualquier tipo de disfunción visual, refractiva, acomodativa y/o binocular, puede provocar la aparición de síntomas visuales que afecten al nivel de confort del paciente cuando se realiza una tarea visual. Aunque estos síntomas son frecuentes, la literatura científica muestra disparidad en cuanto a cuáles son los que están asociados a las disfunciones, tanto en el modo de preguntar por ellos como en la forma de calibrar su severidad. En lo que se refiere a la recogida de los síntomas, existen diversos cuestionarios que evalúan la sintomatología visual en determinadas condiciones, como el uso de las lentes de contacto, el síndrome visual informático, el disconfort visual o la insuficiencia de convergencia. Sin embargo, hasta la fecha no se ha desarrollado ningún cuestionario que analice los síntomas de cualquier disfunción visual (refractiva, acomodativa y/o binocular), a pesar de que sería interesante disponer de una herramienta validada con la que poder ayudar en el diagnóstico de estas anomalías y monitorizar su tratamiento. Es por ello que el objetivo general de esta investigación es desarrollar y validar un cuestionario sobre la sintomatología asociada a las disfunciones visuales, que pueda ser utilizado en una población clínica.

Para el desarrollo y validación de un cuestionario no existen unas pautas definidas ni unos pasos estrictos a seguir, pero sí se recomiendan varios tipos de análisis a incluir para que el cuestionario tenga una buena calidad, y estos deben ser: la identificación de los ítems, la selección de los ítems, el análisis psicométrico, la exactitud diagnóstica, la validez y la repetibilidad.

En esta tesis doctoral se realizaron todos los análisis reseñados. En primer lugar, para identificar los ítems, se analizó la opinión de los profesionales de la visión mediante una metodología Delphi, preguntando sobre qué síntomas estaban asociados a las

disfunciones visuales y cuáles de ellos deberían utilizarse en un cuestionario sobre sintomatología visual. Para el desarrollo de la metodología Delphi, se constituyó un grupo coordinador y un grupo de 14 expertos optometristas españoles de reconocido prestigio profesional. El grupo coordinador planteó al grupo de expertos el problema en cuestión, tratándolo en tres rondas.

Partiendo de los resultados obtenidos en metodología Delphi, para la selección de ítems se desarrolló un cuestionario con 47 preguntas basadas en los síntomas visuales recomendados por los expertos participantes del método Delphi. Se realizó una primera reducción cualitativa mediante un pretest cognitivo, eliminando los ítems redundantes y quedando el cuestionario reducido a 33 ítems. Este cuestionario piloto, denominado Symptom Questionnaire for Visual Dysfunctions (SQVD), se pasó a una muestra de 125 sujetos de un establecimiento sanitario de óptica. Se incluyeron a sujetos con edades entre 15 y 84 años y ausencia de cualquier patología ocular. Se realizó una reducción psicométrica de ítems mediante análisis Rasch, incluyendo en este análisis las curvas de probabilidad de respuesta, ajuste de los ítems al modelo usando los estadísticos infit y outfit, dependencia local a través del estadístico Yen-Q3, análisis del funcionamiento diferencial de los ítems (differential item functioning, DIF), fiabilidad de la separación de las medidas de las personas y de los ítems, unidimensionalidad mediante el análisis de componentes principales de los residuales y medida del alineamiento entre ítems y personas (targeting). Utilizando el análisis Rasch también se calculó la tabla de transformación de puntuación ordinal a intervalos. Tras realizar todos estos análisis, se obtuvo una versión del SQVD de 14 ítems con tres categorías de respuesta, donde los datos se ajustan adecuadamente al modelo de Rasch.

Con el fin de validar el cuestionario SQVD de 14 ítems, éste se distribuyó en una muestra clínica de 306 pacientes de una consulta optométrica, con edades entre 14 y

87 años, donde 204 sujetos tenían algún tipo de disfunción visual (132 disfunciones refractivas y 72 disfunciones de la visión binocular). Se realizó un análisis Rasch para evaluar las propiedades psicométricas del cuestionario donde se evaluaron los siguientes parámetros: curvas de probabilidad de respuesta, ajuste de los ítems al modelo usando los estadísticos infit y outfit, funcionamiento diferencial de los ítems (DIF), fiabilidad de la separación de las medidas de las personas y de los ítems, medida del alineamiento entre ítems y personas (targeting), dependencia local a través del estadístico Yen-Q3 y unidimensionalidad mediante el análisis de componentes principales de los residuales. Además, se analizó la exactitud diagnóstica mediante las curvas ROC (Receiver Operator Characteristic) utilizando como Gold-estándar los síntomas descritos en la historia clínica. También se evaluó la validez del cuestionario mediante la validez concurrente, medida que representa el nivel de correlación entre la puntuación del cuestionario y la puntuación de las medidas clínicas, y la validez de grupos conocidos, análisis que mide el grado en que el instrumento puede discriminar entre grupos clínicamente diferentes. Por último, se examinó la repetibilidad del cuestionario mediante el coeficiente de correlación intraclase (ICC), distribuyéndolo de nuevo a 50 personas una semana después de la primera visita. Los resultados del análisis psicométrico mostraron que el SQVD tiene categorías de respuestas ordenadas, los 14 ítems se ajustaron al modelo, no existiendo DIF significativo para sexo, presbicia y disfunciones, la fiabilidad de la separación de las personas e ítems fue adecuada, al igual que el targeting, además de mostrar unidimensionalidad y no presentar dependencia local entre los ítems. Asimismo, el SQVD obtuvo buena exactitud diagnóstica, adecuada validez concurrente, buena validez de grupos conocidos y muy buena repetibilidad.

Como conclusión, los resultados de esta tesis doctoral muestran que el SQVD tiene buenas propiedades psicométricas, presenta buena exactitud diagnóstica, siendo

un cuestionario válido y repetible para detectar la presencia y frecuencia de síntomas visuales relacionados con cualquier tipo de disfunción visual, bien sea refractiva, acomodativa o binocular. Por lo tanto, esta herramienta puede usarse en la práctica clínica para identificar sujetos con síntomas asociados a las anomalías visuales, así como en estudios de investigación.



Universitat d'Alacant
Universidad de Alicante

1.1. ANTECEDENTES Y ESTADO ACTUAL DEL TEMA

1.1.1. Introducción

En la actualidad, es común encontrar en la consulta optométrica diaria a personas que acuden refiriendo síntomas visuales. Esta sintomatología suele estar generada por las disfunciones visuales, pudiendo ser una disfunción refractiva, acomodativa y/o binocular. Actualmente, se utilizan cada vez más dispositivos electrónicos (ordenadores, tablets, teléfonos móviles, etc.), y este tipo de actividades puede dar lugar a que el sistema visual no desarrolle la tarea con total eficacia debido a que exista alguna disfunción visual. Es bastante común encontrar personas aquejadas de visión borrosa, dolor de cabeza o problemas de enfoque. Este tipo de síntomas pueden disminuir el rendimiento visual, sobre todo en la realización de tareas en visión próxima. De hecho, la mayoría de las tareas de la vida cotidiana se desarrollan en visión próxima, ya sea por estudiar en el caso de los escolares o universitarios, o por trabajar, utilizando dispositivos electrónicos, por lo que la sintomatología que provocan las

disfunciones visuales (refractivas, acomodativas y binoculares) son comúnmente referidas en las personas que acuden al optometrista.

Las disfunciones refractivas, como la miopía, la hipermetropía, el astigmatismo y la presbicia,¹⁻⁴ suelen provocar una serie de síntomas. Habitualmente, una persona que acude a la consulta optométrica lo hace porque considera que no ve bien en alguna distancia (lejos, cerca o ambas). Aunque, a veces, también puede acudir porque, independientemente de ver bien o no, refiere otro tipo de síntomas como dolores de cabeza, disconfort visual, cansancio visual, etc. El hecho de llevar o no la corrección refractiva adecuada puede influir en la aparición de síntomas relacionados con la visión, de ahí la importancia de que las personas lleven la refracción correcta. Además, aunque una persona lleve la prescripción adecuada, pueden existir otro tipo de disfunciones, las acomodativas y las binoculares, anomalías que pueden provocar una sintomatología importante aun llevando corregido el error refractivo.⁵ Estas disfunciones, acomodativas y binoculares, se producen cuando existe un problema en el sistema acomodativo y/o vergencial.

Como se ha dicho anteriormente, si en los anteriores sistemas, refractivo, acomodativo o binocular, aparece alguna anomalía, aparecerán las denominadas disfunciones visuales,⁵ provocando sintomatología y disminuyendo el rendimiento visual de la persona.⁶ De este modo, para su diagnóstico es necesario la realización de un adecuado examen optométrico que incluya no sólo un examen refractivo, sino también pruebas de la visión binocular. Para el examen refractivo debe realizarse en primer lugar la toma de la agudeza visual (AV),² la retinoscopía,^{2,7} y partiendo de este valor se realiza el examen subjetivo hasta alcanzar el máximo positivo con la máxima agudeza visual,⁸ afinando el eje y el cilindro mediante los cilindros cruzados de Jackson.⁸ Una vez que se obtiene el máximo positivo con la máxima agudeza visual,³ este resultado se debe utilizar como base para todas las pruebas de la visión binocular.

Dentro del examen de la visión binocular deben realizarse pruebas acomodativas y binoculares. Las pruebas acomodativas incluyen la medida de la amplitud de acomodación, la flexibilidad acomodativa, la respuesta acomodativa y las acomodaciones relativas.

La amplitud de acomodación (AA)⁹ determina la capacidad máxima de acomodación de una persona que es necesaria para mantener nítida la imagen de un objeto. De manera monocular, se le acerca lentamente al paciente un optotipo de AV 20/20 hasta que el sujeto refiera borrosidad mantenida en el test.¹⁰ La posición final del optotipo en centímetros (cm) se expresa en dioptrías (D) (1/cm), obteniendo así la medida de la AA.^{11,12}

La flexibilidad acomodativa (FA) es la capacidad del sistema acomodativo para realizar cambios rápidos de acomodación,⁹ pudiendo realizarse de modo monocular (FAM) o binocular (FAB). Para su medida¹³ se utiliza un flipper de ± 2.00 D y un optotipo de AV 20/30, con control anti supresión para medir la FAB. La prueba se realiza durante un minuto y se contabilizan los ciclos por minuto (cpm), considerando un ciclo el cambio desde una lente positiva a otra.

El método de estimación monocular (MEM) es la medida de la respuesta acomodativa mediante retinoscopia dinámica.¹⁴ En esta prueba se utiliza un test a 40 centímetros de agudeza visual 20/30¹⁵ utilizando gafas de prueba con lentes de prueba.

Para evaluar la capacidad de variación del sistema acomodativo manteniendo constante el estímulo vergencial se utilizan las acomodaciones relativas.⁹ La acomodación relativa positiva (ARP) y negativa (ARN) se miden con lentes negativas y positivas respectivamente, utilizando un optotipo acomodativo de AV 20/30 a 40 centímetros. Dichas lentes se van insertando hasta que el paciente refiere borrosidad.^{16,17}

Dentro de las pruebas binoculares, se incluyen el cover test, el cociente AC/A, el punto próximo de convergencia, vergencias fusionales, flexibilidad vergencial, disparidad de fijación, fusión y estereopsis.

El cover test es una de las pruebas binoculares más importantes pues determina la presencia, dirección y magnitud de cualquier tipo de desviación de los ejes visuales.⁹ Consta de dos partes, el cover unilateral que sirve para diagnosticar una tropia, y el cover alternante para determinar el valor de la desviación.¹⁸ Ambas medidas se realizan en visión lejana y en próxima mientras el sujeto fija una letra de AV 20/30,^{5,18-21} donde es de especial importancia que la oclusión sea adecuada, durante al menos 5 segundos.²²⁻²⁵ Para su medida se utiliza una barra de prismas mientras se realiza el cover alternante, utilizando base nasal (BN) para exoforia y base temporal (BT) para endoforia, obteniendo la desviación en dioptrías prismáticas (Δ). La barra prismática debe situarse en posición plano frontal, a una distancia no más lejana de 1 cm desde el plano de las gafas.^{18,21} El valor de la desviación se obtiene anotando el punto medio entre el primer valor prismático con el que no se observa movimiento y el valor que provoca la inversión.²⁶

El cociente AC/A se describe como la cantidad de vergencia acomodativa que se puede estimular por unidad de acomodación.²⁷ Para determinar el AC/A mediante el método gradiente se obtiene la desviación en visión próxima y a continuación se repite esta misma medida de la desviación, pero añadiendo una lente de -1.00 D. La diferencia entre los valores de las desviaciones es el valor AC/A gradiente.¹⁷

El punto próximo de convergencia (PPC) es la capacidad máxima de convergencia que tiene una persona.⁹ La medida del PPC se realiza aproximando un optotipo de AV 20/30^{21,28-30} al paciente hasta que éste observa la diplopía (punto de ruptura) y a continuación se aleja hasta que recobra la fusión (punto de recobro).³¹⁻³⁴

Las vergencias fusionales positivas (VFP) y negativas (VFN) son los movimientos fusionales que se producen por un cambio en la disparidad retiniana.⁵ Dicha medida se puede realizar mediante los prismas de Risley del foróptero, mientras el sujeto fija un optotipo de AV 20/30^{5,16,35} tanto en visión lejana como próxima. Para medir las VFP se insertan prismas de base temporal, y para las VFN de base nasal,³⁶ hasta que el paciente refiere borrosidad (punto de borrosidad) y luego doble (punto de ruptura). En ese momento, se disminuye la potencia prismática hasta que el paciente refiere una imagen simple (punto de recobro).⁹

La flexibilidad vergencial es la habilidad del sistema vergencial para producir cambios vergenciales rápidos.³⁷ Para medir la flexibilidad vergencial, se utiliza un optotipo de AV 20/30^{38,39} para 40 centímetros, introduciendo una combinación prismática de 12BT/3BN.³⁹

La disparidad de fijación (DF) representa el error de vergencia necesario para mantener activo el sistema vergencial. La valoración de la DF se realiza en condiciones de visión binocular,⁹ pudiendo existir endodisparidad de fijación (EDF) o exodisparidad de fijación (XDF).

La fusión representa la interpretación de dos imágenes similares en la percepción única, evaluando la presencia y el tamaño de un escotoma de supresión.⁵ Se evalúa mediante la linterna de Worth, observando si el paciente fusiona o presenta supresión desde su determinación en visión próxima hasta visión lejana.⁵

Finalmente, la medida de la estereopsis es una prueba binocular con la que se evalúa el grado máximo de binocularidad, es decir, la mínima disparidad binocular que puede percibir un sujeto medida en segundos de arco (").⁴⁰

En la Tabla 1 se pueden ver un resumen de los valores normativos, basados en una recopilación de estudios realizada por Cacho et al,⁹ y utilizados para las pruebas descritas, con los valores esperados y los resultados con los que se considera que el sujeto ha fallado la prueba, valor necesario para conocer cuándo su resultado se encuentra fuera de los rangos esperados.



Universitat d'Alacant
Universidad de Alicante

PRUEBA	VALORES ESPERADOS (media ± DE)	FALLA LA PRUEBA
AA	AA mínima 15 - 0.25 Edad	2 D < AA mínima
FAM	11 ± 5 cpm	< 6 cpm
FAB	8 ± 5 cpm	< 3 cpm
MEM	0.50 ± 0.25 D	≤ 0 D / ≥ 1 D
ARP	2.37 ± 1.12 D	≤ 1.00 D
ARN	2.00 ± 0.50 D	≤ 1.25 D
Cover Test lejos / cerca	1 X (exo) ± 2 Δ / 3 X' (exo) ± 5 Δ	La mitad de la DE Lejos: ≥ 1E / ≥ 3 X Cerca: ≥ 1 E' / ≥ 6X'
AC/A	4/1 ± 2 Δ/D	≤ 1/1 / ≥ 7/1 Δ/D
PPC	2.5 ± 2.5 cm / 4.5 ± 3 cm	2 veces la DE ≥ 7.5/10.5 cm
VFP suaves (lejos)	9 ± 4 Δ / 19 ± 8 Δ / 10 ± 4 Δ	≤ 4/10/5 Δ
VFN suaves (lejos)	x / 7 ± 3 Δ / 4 ± 2 Δ	≤ X/3/1 Δ
VFP suaves (cerca)	17 ± 5 Δ / 21 ± 6 Δ / 11 ± 7 Δ	≤ 11/14/3 Δ
VFN suaves (cerca)	13 ± 4 Δ / 21 ± 4 Δ / 13 ± 5 Δ	≤ 8/16/7 Δ
VF vertical	3 ± 1 Δ / 2 ± 1 Δ	≤ 1/0 Δ
Flexibilidad de vergencia	16 ± 2.6 cpm	< 13.4 cpm
DF	Hasta 6' XDF / 4' EDF	> 6' XDF / 4' EDF
Fusión y estereopsis	30 - 50"	Supresión / > 50"

Tabla 1. Valores normativos para las pruebas acomodativas y binoculares basados en una recopilación de estudios realizada por Cacho et al.⁹ (DE: Desviación estándar; AA: Amplitud de acomodación; FAM: Flexibilidad acomodativa monocular; FAB: Flexibilidad acomodativa binocular; MEM: Método de estimación monocular; ARP: Acomodación relativa positiva; ARN: Acomodación relativa negativa; PPC: Punto próximo de convergencia; VFP: Vergencias fusionales positivas; VFN: Vergencias fusionales negativas; VF: Vergencia fusional; DF: Disparidad de fijación; D: dioptrías; Δ: Dioptría prismática; cpm: ciclos por minuto; E: endoforia en lejos; E': Endoforia en cerca; X: exoforia en lejos; X': Exoforia en cerca EDF: Endodisparidad de fijación; XDF: Exodisparidad de fijación; “: Segundos de arco; cm: centímetros).

1.1.2. Disfunciones Visuales

Para el diagnóstico de las disfunciones visuales, por un lado, la información obtenida en la retinoscopia y el examen subjetivo será la utilizada para diagnosticar las disfunciones refractivas. En ocasiones el sujeto tendrá un error refractivo mal corregido que necesite ser modificado o en otras puede existir un error refractivo no corregido. Estos errores refractivos, como ya se conoce, son miopía, hipermetropía, astigmatismo y presbicia.¹⁻⁴

Por otro lado, para el diagnóstico de las disfunciones de la visión binocular, será necesario además del examen refractivo, utilizar las pruebas acomodativas y binoculares comentadas anteriormente.

En cuanto a la clasificación de las anomalías acomodativas se incluyen^{5,41,42} la insuficiencia acomodativa (IA), exceso acomodativo (EA) e inflexibilidad acomodativa. Las disfunciones binoculares no estrábicas se clasifican según la clasificación realizada por Wick⁴³ y utilizada actualmente,^{5,41} y en ella se tiene en cuenta la relación entre la heteroforia de lejos y de cerca y el cociente AC/A, dividiéndose de la siguiente manera: disfunciones de AC/A bajo (insuficiencia de convergencia (IC) e insuficiencia de divergencia (ID)), disfunciones de AC/A alto (exceso de convergencia (EC) y exceso de divergencia (ED)), disfunciones de AC/A normal (disfunción de vergencia fusional (DVF), endoforia básica (EB) y exoforia básica (XB)). Dentro de las disfunciones binoculares también se incluyen las desviaciones verticales. En cuanto a las disfunciones binoculares estrábicas, su clasificación en endotropía y exotropía es más compleja puesto que las desviaciones pueden ser concomitantes o inconcomitantes.⁴⁴

Para diagnosticar las disfunciones acomodativas y binoculares no estrábicas, en general, se suelen usar las pautas descritas en la literatura científica teniendo en cuenta los signos clínicos derivados de estas pruebas acomodativas y binoculares.^{5,9,42,45-48} En

este sentido, Cacho-Martínez et al⁴² muestran en una revisión sistemática que existen pocos estudios que analicen la exactitud diagnóstica de las pruebas que deben utilizarse para su diagnóstico, además de existir gran disparidad de criterios diagnósticos entre los distintos autores tanto para las disfunciones acomodativas como para las binoculares. En esta revisión sistemática⁴² se demuestra que, para las condiciones binoculares, los estudios existentes utilizan diferentes signos clínicos, no existiendo investigaciones que hayan evaluado la exactitud diagnóstica de los signos clínicos que deben utilizarse para el diagnóstico de las anomalías binoculares. Los estudios no analizan los valores predictivos, la sensibilidad y especificidad, o el análisis de las curvas ROC (Receiver Operating Characteristic, ROC). Sin embargo, esta revisión sistemática⁴² sí que demuestra que hay estudios de validez diagnóstica para las disfunciones acomodativas. De este modo, los resultados de esta revisión sistemática⁴² y posteriores estudios de validez diagnóstica,⁴⁵⁻⁴⁷ permiten establecer un diagnóstico de estas anomalías de acuerdo con el número de signos clínicos asociados a cada disfunción, clasificando estos signos como fundamentales y complementarios. Por lo tanto, basándose en los estudios que refieren la validez diagnóstica de los signos clínicos,^{9,42,45-48} en la Tabla 2 se documentan cuáles son los signos fundamentales y complementarios para el diagnóstico de las disfunciones acomodativas y binoculares no estrábicas.

En el caso de las ambliopías y estrabismos, el análisis de la agudeza visual y el cover test será necesario para su diagnóstico. En el caso de las tropías, el examen optométrico debe incluir pruebas específicas que permitan evaluar la desviación y la concomitancia.⁴⁴

DISFUNCIÓN	SIGNOS FUNDAMENTALES	SIGNOS COMPLEMENTARIOS
Insuficiencia de acomodación	AA disminuida	FAM falla (-), FAB falla (-), MEM alto, ARP baja
Exceso de acomodación	FAM falla (+)	FAB falla (+), MEM bajo, ARP alta, ARN baja
Inflexibilidad acomodativa	FAM falla (+) y (-)	FAB falla (+) y (-) ARP baja, ARN baja
Insuficiencia de convergencia	Exoforia en cerca significativa > lejos, PPC alejado, FAB falla (+)	VFP reducida en cerca, FV falla con 12 Δ, MEM bajo, ARN alta
Insuficiencia de divergencia	Endoforia en lejos significativa > cerca	VFN reducida en lejos y cerca, FV falla con 3 Δ, FAB falla (-), MEM alto, ARP baja
Exceso de convergencia	Endoforia en cerca significativa > lejos, FV falla con 3 Δ, FAB falla (-)	VFN reducida en cerca, MEM alto, ARP baja
Exceso de divergencia	Exoforia en lejos significativa > cerca	VFP reducida en lejos y cerca, PPC alejado, FV falla con 12 Δ, FAB falla (+), MEM bajo, ARN baja
Disfunción de vergencia fusional	VFP y VFN reducidas en lejos y cerca o FV falla con 3 Δ y 12 Δ	FAB falla con (+) y (-), ARP baja, ARN alta
Endoforia básica	Endoforia significativa en lejos y cerca de aprox. el mismo valor	VFN reducida en lejos y cerca, FV falla con 3 Δ, FAB falla (-), MEM alto, ARP baja
Exoforia básica	Exoforia significativa en lejos y cerca de aprox. el mismo valor	VFP reducida en lejos y cerca, PPC alejado, FV falla con 12 Δ, FAB falla (+), MEM bajo, ARN baja
Disfunción vertical	Hiperforia $\geq 1 \Delta$	Vergencias verticales reducidas, VFP y VFN reducidas en lejos y cerca, FV falla con 3 Δ y 12 Δ, anomalías en la posición de la cabeza

Tabla 2. Signos fundamentales y complementarios para el diagnóstico de las disfunciones acomodativas y binoculares no estrábicas, descritos por García-Muñoz et al.⁴⁹ (AA: Amplitud de acomodación; FAM: Flexibilidad acomodativa monocular; FAB Flexibilidad acomodativa binocular; MEM: Método de estimación monocular; ARP: Acomodación relativa positiva; ARN: Acomodación relativa negativa; PPC: Punto próximo de convergencia; VFN: Vergencias fusionales negativas; VFP: Vergencias fusionales positivas; FV: Flexibilidad vergencial; Δ: Dioptría prismática).

En lo que se refiere a la prevalencia de las disfunciones visuales, para los errores refractivos los valores de prevalencia global⁵⁰ oscilan entre 11.4% y 44.8%. En el caso de las disfunciones acomodativas y binoculares no estrábicas, existe disparidad en la literatura científica sobre sus valores de prevalencia, variando entre 0.2% y 61.7%^{49,51-63} para las anomalías acomodativas y en el caso de las disfunciones binoculares no estrábicas, entre un 0.1% y un 33%.^{49,51-54,56,57,59-62,64-69} La Tabla 3 especifica los valores de prevalencia para cada una de las disfunciones específicas. En cuanto a la ambliopía, la estimación de la prevalencia global es de 1.44%, siendo de 1.94% para los menores de 5 años y de 3.29% para los mayores de 20 años.⁷⁰ Y en el caso del estrabismo, la estimación de la prevalencia global es de 1.93%, siendo de 1.78% para menores de 20 años y 3.29% para mayores de 20 años.⁷¹

Una vez se han diagnosticado, las disfunciones visuales requieren un tratamiento para mejorar los síntomas y signos. En las disfunciones refractivas, el tratamiento de los sujetos es prescribir una corrección óptica. En el caso de las disfunciones acomodativas y binoculares no estrábicas existen diferentes pautas en el tratamiento referidas por la literatura científica, como pueden ser la terapia visual, la prescripción de prismas o la prescripción de adiciones positivas.^{54,55,57,72-99} Asimismo, en el caso de las ambliopías el tratamiento puede ser desde la prescripción del error refractivo, la oclusión, la terapia visual y la atropina.¹⁰⁰⁻¹⁰⁷ Para los estrabismos, el tratamiento puede ser cirugía, prescripción de prismas, prescripción del error refractivo, oclusiones y terapia visual.^{44,108-110}

	DISFUNCIÓN	PREVALENCIA (%)
Disfunciones refractivas	Miopía	16.2 – 32.9 ⁵⁰
	Hipermetropía	23.1 – 38.6 ⁵⁰
	Astigmatismo	11.4 – 44.8 ⁵⁰
Disfunciones acomodativas y binoculares no estrábicas	Insuficiencia de acomodación	0.2 – 61.7 ^{49,51-63}
	Exceso de acomodación	0.8 – 10.8 ^{49,51-53,55,56}
	Inflexibilidad acomodativa	0.4 – 33 ^{49,52,53,55,56,60}
	Insuficiencia de convergencia	2 – 33 ^{49,51-54,56,57,59-62,64-68}
	Insuficiencia de divergencia	0.1 – 7 ^{51,56,60,66,68}
	Exceso de convergencia	0.8 – 15 ^{49,51-54,56,60,61,66,68}
	Exceso de divergencia	0.4 – 1.1 ^{51,56,57,60,66,68}
	Disfunción de vergencia fusional	0.4 – 3.3 ^{49,51,54,56,57,60,66}
	Endoforia básica	0.3 – 9 ^{49,51,56,60,66,68}
	Exoforia básica	0.3 – 10.8 ^{49,51-53,60,66,68}
	Disfunción vertical	0.2 ⁶⁹

Tabla 3. Prevalencia de las disfunciones visuales.

Universitat d'Alacant
 Universidad de Alicante

1.1.3. Síntomas visuales

Las disfunciones visuales descritas anteriormente son comunes en la práctica clínica habitual y, como se ha comentado, pueden provocar una serie de síntomas visuales.¹¹¹ Uno de los principales problemas que le surgen a los optometristas cuando se realiza un examen optométrico, es determinar la existencia real de síntomas con los que se asocian las disfunciones visuales. Tampoco existe un consenso sobre los síntomas específicos de cada una de las disfunciones visuales,¹¹² ya que muchos de estos síntomas son compartidos entre estas anomalías visuales.

Por un lado, en la literatura científica se refiere que las disfunciones refractivas suelen presentar síntomas variados, como astenopía, cansancio ocular, diplopía, disconfort visual, dolor de cabeza, dolor ocular, entornar los ojos, fatiga visual, ojo rojo, ojos cansados, ojos irritados, ojos llorosos, ojos secos o arenosos y visión borrosa.^{111,113}

En cuanto a las disfunciones de la visión binocular, los síntomas también son variados. García-Muñoz et al¹¹² encontraron en una revisión sistemática al respecto, que existían 267 formas diferentes de expresar los síntomas en 56 estudios de la literatura científica, lo que indica que existe una gran variabilidad respecto a qué síntomas se asocian a estas anomalías. Después de un proceso cualitativo de categorización, los autores agruparon estos síntomas en 34 categorías (Tabla 4), que fundamentalmente se relacionaban con la visión próxima. Los síntomas más comunes hallados en esta revisión¹¹² para las disfunciones de la visión binocular fueron dolor de cabeza, visión borrosa, diplopía, fatiga visual y movimiento o parpadeo de las palabras en la visión de cerca.

Síntomas
Astenopía
Calcular mal las distancias
Dar sueño
Dificultad para cambios de enfoque
Dificultad para realizar las tareas escolares
Diplopía
Discomfort visual
Dolor de cabeza
Dolor de estómago o náusea
Dolor ocular
Evitar realizar tareas
Excesiva sensibilidad a la luz
Falta de concentración
Fatiga visual
Frotarse los ojos
Inclinar el libro o la cabeza
Lagrimo
Modificar la distancia de trabajo
Movimiento o parpadeo del texto en cerca
Notar que el ojo se desvía
Ocluir un ojo
Ojo rojo
Ojos irritados
Ojos llorosos
Ojos secos o arenosos
Palabras con halos de colores
Parpadeo ocular excesivo
Perderse al leer
Pesadez ocular
Postura inadecuada en tareas de cerca
Problemas de lectura
Problemas en el aprendizaje
Ser distraído
Visión borrosa

Tabla 4. Categorías de síntomas localizadas en la revisión sistemática de García-Muñoz et al sobre las diferentes disfunciones de la visión binocular.¹¹²

Además, en esa investigación¹¹² se observó que 15 de estas 34 categorías de síntomas eran comunes para las disfunciones acomodativas y binoculares. Estos síntomas fueron:

- astenopía
- dificultad para realizar tareas escolares
- diplopía
- disconfort visual
- dolor de cabeza
- dolor ocular
- evitar realizar tareas
- falta de concentración
- fatiga visual
- modificar la distancia de trabajo
- movimiento o parpadeo del texto en cerca
- ojos irritados
- problemas de lectura
- problemas en el aprendizaje
- visión borrosa

Todos estos síntomas suelen estar clasificados en la literatura científica y optométrica por su asociación a la distancia de trabajo, donde en su mayoría se asocian a la visión próxima.^{5,112}

Por otra parte, en otra investigación¹¹¹ se caracterizó la sintomatología asociada a las disfunciones visuales en una población universitaria, basándose en las 34 categorías de síntomas de la Tabla 4. En este estudio¹¹¹ se caracterizó la sintomatología recogida en la anamnesis del examen optométrico, y se encontró que una disfunción refractiva no corregida puede contaminar la sintomatología producida por las anomalías acomodativas y binoculares, particularmente en las disfunciones acomodativas. Esto significa que, cuando un paciente presente una disfunción acomodativa además de un error refractivo no corregido, no se puede asegurar que sus síntomas sean debidos a dicha disfunción acomodativa, por lo que primero se debe prescribir el error refractivo antes de iniciar un tratamiento específico para la disfunción acomodativa. Además, en

esta misma investigación se encontró una asociación entre la sintomatología visual y las disfunciones binoculares, de modo que cuanto mayor es el número de signos clínicos utilizados para el diagnóstico de la disfunción binocular, mayor es la probabilidad de presentar sintomatología.

En resumen, como se puede observar, no existe un consenso sobre qué síntomas están relacionados específicamente con cada una de estas anomalías, observando que pocos síntomas son específicos de cada disfunción y muchos de ellos son compartidos entre disfunciones.

1.1.4. Herramientas para la recogida de síntomas visuales

Además de la existencia de esta variedad en cuanto a los síntomas relacionados con las disfunciones visuales, también existe diversidad en la manera de obtenerlos. La forma más común que tienen los optometristas de recopilar esta sintomatología es mediante la historia clínica, aunque algunas veces también se realiza mediante cuestionarios. Cuando se utiliza la historia clínica, los optometristas suelen, en algunas ocasiones, anotar la descripción de los síntomas tal y como los pacientes los comentan, y en otras, la sintomatología se refleja en función de las preguntas que plantea el profesional que realiza el examen optométrico. Estas situaciones producen diferencias a la hora de nombrar los síntomas y especificar su frecuencia.

También ocurre lo mismo cuando se utilizan cuestionarios. Existen diferentes formas de preguntar por un mismo síntoma entre diferentes cuestionarios. A su vez, existen diferentes escalas de medición para medir la sintomatología, donde algunos cuestionarios utilizan escalas dicotómicas (dos respuestas) o politómicas (tres o más respuestas). Dependiendo del contexto y de la condición visual a la que va referido el

síntoma, los cuestionarios pueden adaptar la forma de preguntar el síntoma y la escala adecuada para recoger su respuesta.

En este sentido, se han desarrollado instrumentos para algunas condiciones visuales, como cuestionarios para baja visión,¹¹⁴ cataratas,^{115,116} queratocono,¹¹⁷ usuarios de lentes de contacto,¹¹⁸ síntomas asociados a vídeo terminales o pantallas de ordenador,^{119,120} instrumentos relacionados con la calidad de la visión,¹²¹ calidad de vida relacionada con la visión¹²²⁻¹²⁶ o desviaciones verticales.¹²⁷

En lo que se refiere a cuestionarios específicos para disfunciones visuales, los resultados de la revisión sistemática de García-Muñoz et al¹¹² muestran la existencia de 11 cuestionarios diferentes para la recogida de los síntomas.^{34,128-138} Algunos utilizan una escala de respuesta dicotómica y otros utilizan una escala de opciones de respuesta múltiple de hasta 5 respuestas por ítem. De estos 11 cuestionarios,^{34,128-138} solamente tres están validados, refiriéndose dos de ellos a la insuficiencia de convergencia, el Convergence Insufficiency Symptom Survey (CISS) versión para niños y versión para los padres,^{34,131,132} y el cuestionario de Conlon,¹³⁰ cuestionario relacionado con el discomfort visual y no asociándose a ninguna disfunción. Como se observa, no existe ningún cuestionario validado de sintomatología visual asociada a cualquier tipo de disfunción visual, por lo que sería interesante desarrollar y disponer de una herramienta validada con la que poder ayudar en el diagnóstico de estas anomalías y monitorizar su tratamiento, además de poder utilizarlo en estudios de investigación.

1.1.5. Desarrollo y validación de cuestionarios

Para el desarrollo y validación de un cuestionario se deben seguir una serie de recomendaciones. No existen unas pautas definidas ni unos pasos estrictos a seguir,

pero sí que se recomienda a la hora de desarrollarlo, tener en cuenta varios tipos de análisis a incluir en este proceso para que el cuestionario tenga una calidad alta. En los últimos años han aparecido en la literatura científica varias guías con recomendaciones para el desarrollo y validación de cuestionarios.¹³⁹⁻¹⁴¹ Considerando estas recomendaciones, para poder desarrollar y validar un cuestionario se deberían realizar los siguientes análisis: identificación de los ítems, selección de los ítems, análisis psicométrico, exactitud diagnóstica, validez y repetibilidad.

1.1.5.1. Identificación de los ítems

El primer paso para la elaboración de un cuestionario es identificar qué ítems se deben incluir en éste. Para ello, puede ser muy importante conocer la opinión de los profesionales de la visión sobre qué síntomas clínicos considerarían que están relacionados con las anomalías visuales. Para poder conocer esa opinión, en la atención sanitaria es habitual utilizar métodos de investigación cualitativa. Dentro de estos métodos investigación cualitativa existen varios tipos de métodos de consenso, divididos en dos grupos: los grupos focales y los métodos formales de consenso.

Los grupos focales son útiles para investigar la identificación de problemas, preguntas o cuestiones importantes.¹⁴² En los grupos focales existe la figura del moderador que se encarga de hacer preguntas a los expertos y dirigir la discusión para no alejarse del tema (de ahí su nombre), y solo requieren alrededor de varias horas para completar su tarea.¹⁴² Por otro lado, dentro de los métodos formales de consenso, existen varias técnicas: técnica de grupo nominal,¹⁴³ conferencia de consenso,¹⁴⁴ método Delphi^{145,146} y el método Rand/UCLA.¹⁴⁷

La técnica de grupo nominal (también conocida como panel de expertos) es un proceso en el que se reúne un grupo de expertos para identificar de un problema, generar soluciones y tomar decisiones.¹⁴³ Con ayuda de un moderador, el grupo de expertos debe discutir el problema planteado, y tras haber elaborado un listado de ideas, estas ideas se deben votar para luego jerarquizar los resultados y así poder tener en cuenta las opiniones de todo el mundo.

La conferencia de consenso es una reunión abierta de expertos que se reúnen para llegar a un consenso sobre un tema, complementado con unos expertos que no toman parte en las decisiones del grupo.¹⁴⁴ Tras escuchar a este último grupo de expertos que no toma las decisiones, el primer grupo de expertos deben llegar a un consenso con las opiniones del segundo grupo.

El método de consenso formal más utilizado en el ámbito sanitario es el método Delphi,^{145,146} método considerado como una forma eficiente de recopilar información de un grupo de expertos.¹⁴⁵ El método Delphi se define como un “método de estructuración de un proceso de comunicación grupal que es efectivo a la hora de permitir a un grupo de individuos, como un todo, tratar un problema complejo”.¹⁴⁸⁻¹⁵⁰ Es una técnica de comunicación prospectiva estructurada, fundamentalmente para obtener información cualitativa tomando en consideración la opinión de cada miembro que conforma el panel de expertos. El método Delphi permite a los expertos comunicar sus opiniones y conocimientos sobre un problema complejo con el fin de explorar opciones y llegar a un consenso, incluso si se encuentran en áreas geográficamente dispersas.¹⁵¹ De hecho, se ha demostrado que el método Delphi es una herramienta útil para llegar a un consenso en un área de incertidumbre o falta de evidencia empírica.¹⁵² Este método tiene varias características que lo diferencian de otros métodos de consenso:^{145,153,154} anonimato de la respuesta (los miembros no se conocen, lo que les permite no inhibirse en sus respuestas), interacción (el proceso ocurre en rondas en las que las interacciones

entre los miembros del grupo se realizan a través de un cuestionario en lugar de una comunicación cara a cara) y retroalimentación individual controlada (que muestra la distribución de la respuesta del grupo). Además, existe un análisis estadístico que utiliza medidas descriptivas de la respuesta del grupo completo. No existe un método estándar para calcular el número de expertos para la técnica Delphi, pero se ha sugerido una muestra de aproximadamente quince personas,¹⁵⁵ aunque se pueden utilizar paneles más grandes. La selección de los participantes en el proceso se debe hacer atendiendo a las principales características que la literatura científica recomienda para considerar un experto para una técnica Delphi, es decir, según sus conocimientos y su experiencia.¹⁵⁶

El método Rand/UCLA es otro de los métodos de consenso¹⁴⁷ y es una combinación entre la técnica de grupo nominal y el método Delphi. En esta metodología, los investigadores envían la información al grupo de expertos. En una primera ronda, los expertos otorgan una puntuación a las ideas enviadas por los investigadores y les envían los resultados. Los investigadores, en una segunda ronda, vuelven a enviar los resultados a los expertos, y éstos, de forma presencial, debaten de manera estructurada las propuestas que ellos mismos han realizado. En último lugar, cada experto de forma anónima valora las propuestas más interesantes e intentan llegar a un consenso.

1.1.5.2. Selección de los ítems

Una vez se ha conocido la opinión de los expertos, independientemente del método de investigación cualitativa elegido, y se ha recopilado la información necesaria para poder iniciar el desarrollo de un cuestionario, se debe realizar un pretest cognitivo en el que se incluyen ya preguntas sobre la cuestión clínica que se está estudiando. Es una fase importante a incluir en el desarrollo de cualquier cuestionario¹⁵⁷ cuyo propósito

es verificar que el público objetivo comprende las preguntas y las opciones de respuesta propuestas y que sea capaz de responder de manera significativa. Por lo general, cuando se realiza un pretest cognitivo, el participante completará el cuestionario y luego informará sobre cada elemento al investigador para la identificación de problemas, como por ejemplo comprobar si los enunciados son comprensibles, si tiene la extensión adecuada, si existe rechazo a algunas preguntas, si la estructura del cuestionario es lógica y si la duración del cumplimiento es aceptable. Además, se pueden recoger todos los comentarios que puedan hacer los participantes.

Una vez realizado el pretest cognitivo se diseña el cuestionario piloto que tiene que ser analizado mediante algún método psicométrico, como puede ser el análisis factorial, o con el método actualmente recomendado por la literatura científica en cuanto a desarrollo de cuestionarios, el análisis Rasch.^{139,140} Este cuestionario piloto permitirá eliminar elementos no válidos, poco fiables o que discriminan de manera deficiente.

1.1.5.3. Análisis psicométrico

Para realizar un análisis psicométrico de un cuestionario, como se ha dicho anteriormente, el método recomendable que se debe utilizar es el análisis Rasch. Existen varios métodos estadísticos para analizar psicométricamente un cuestionario: la Teoría Clásica de los Test (Classic Test Theory, CTT) y la Teoría de Respuesta al Ítem (Item Response Theory, IRT). Estas teorías sostienen que, en un determinado cuestionario, todos los ítems miden lo mismo. La CTT hace deducciones a partir de la puntuación total del conjunto de ítems evaluados, no siendo posible hacer inferencias fuera de ese conjunto de ítems que conforman un cuestionario. Por otro lado, la IRT en la cual se basa el análisis Rasch, tiene en cuenta la interacción entre los ítems y las personas. De este modo, determina la probabilidad de una persona de responder con

éxito a un ítem, teniendo en cuenta la capacidad de la persona y la dificultad del ítem. Para ello, la IRT convierte las respuestas ordinales de un cuestionario en una medida lineal, además de poder analizar si la transformación de dichas respuestas tiene la calidad necesaria. La fortaleza de la metodología de Rasch es que permite la medición conjunta de personas e ítems en la misma dimensión o constructo.¹⁵⁸ Este método se recomienda para el análisis de cuestionarios,¹³⁹ ya que proporciona una idea de la consistencia interna de la escala y puede relacionar la dificultad del ítem con la capacidad de la persona.¹²¹

El análisis de Rasch es un modelo probabilístico que asume que los ítems varían en dificultad.¹⁵⁹ Estima la dificultad de los ítems (item difficulty) y las habilidades de las personas (person ability) asociándolos en una escala de intervalo lineal común, lo que permite una fácil comparación entre las medidas.¹¹⁷ Transforma las observaciones cualitativas (escala ordinal) en una medida lineal en logits (log-odds unit), unidad de medida del análisis Rasch.¹²² Aunque el modelo probabilístico del análisis Rasch se planteó en su origen para ítems de respuesta dicotómica, se han formulado otros modelos para respuesta poltómica, como el Modelo de Crédito Parcial (Partial Credit Model, PCM) o el Modelo de Escala de Calificación de Andrich (Andrich Rating Scale Model, ARSM). PCM considera unos umbrales de categoría diferentes para cada ítem, es decir, asume que cada ítem está definido por su propia escala de medida, mientras que ARSM asume umbrales de categoría iguales en todos los ítems, o lo que es lo mismo, todos los ítems tienen el mismo formato y usan la misma escala de calificación categórica. Para saber qué modelo escoger, se utiliza la estadística de razón de verosimilitud (likelihood ratio statistic). Si la prueba es significativa ($p < 0.05$) se utiliza el PCM, y si no es significativa ($p > 0.05$), el ARSM.¹⁶⁰

Una vez que se ha elegido el modelo, se deben realizar varios tests estadísticos dentro del análisis Rasch. Estos análisis son:¹⁴¹ curvas de probabilidad de respuesta,

ajuste de los ítems al modelo usando los estadísticos infit y outfit, dependencia local, análisis del funcionamiento diferencial de los ítems (DIF), fiabilidad de la separación de las medidas de las personas y de los ítems, unidimensionalidad, medida del alineamiento entre ítems y personas (targeting) y tabla de transformación de las puntuaciones.¹⁴¹

A. Curvas de probabilidad de respuesta

Para medir la funcionalidad de las categorías de respuesta, el análisis Rasch proporciona información sobre el número adecuado de categorías de respuesta que deben incluirse en el cuestionario. Para analizar si la calibración de las categorías aumenta de manera ordenada, el análisis Rasch utiliza las curvas de probabilidad de respuesta (category probability curves, CPC).¹⁶¹ Estas curvas muestran la probabilidad de que una persona con un valor específico de la habilidad que se está midiendo en el cuestionario (*person ability*), y en relación con la dificultad del ítem (*item difficulty*), seleccione una categoría de respuesta.¹³⁹ Cada categoría de respuesta se representa mediante una curva. El umbral es el punto de intersección entre las categorías de respuesta adyacentes, de modo que revela el punto en el que la probabilidad de elegir cualquiera de las categorías de respuesta es la misma.¹⁶² Si los umbrales están desordenados, es decir, las categorías de respuesta no son equidistantes entre ellas, se debe colapsar la categoría desordenada en una categoría adyacente.^{139,163,164} Para detectar esta situación, se debe realizar la medida de umbral de Andrich de modo que los umbrales deben estar separados al menos 1.4 logits.¹⁶¹

B. Ajuste de los ítems al modelo usando los estadísticos infit y outfit

Para realizar el análisis del ajuste de los ítems al modelo, en el análisis Rasch se utilizan los estadísticos denominados infit y outfit. Estos estadísticos comparan las respuestas previstas por el modelo con las observadas. Infit es un estadístico de ajuste con información ponderada, que indica el ajuste entre el valor esperado y el observado con respecto a los valores promedio. Detecta desajustes en las desviaciones cerca de la zona de medición del ítem, por ejemplo, que el ítem obtenga siempre la misma respuesta o que las personas contesten siempre la misma categoría, y es menos sensible a los valores atípicos (outliers). Outfit es un estadístico de ajuste sensible a los casos atípicos, es decir, más sensible a outliers, por lo que permite detectar respuestas inusuales que ocurren de manera inesperada.

En el análisis Rasch, se calculan los cuadrados medios (mean squares, MNSQ) de infit y outfit para explorar la compatibilidad de los datos con el modelo.^{159,165} Infit MNSQ y outfit MNSQ deben tener un valor cercano a 1 para que se ajusten a un buen modelo. Además, para considerar que existe un adecuado ajuste al modelo Rasch, estos valores deben estar en un rango entre 0.70 y 1.30 logits.¹³⁹ Valores inferiores a 0.70 indican una posible redundancia de ítems y valores superiores a 1.30 sugieren que los ítems pueden medir algo diferente a la escala general. En general, dichos estadísticos infit y outfit MNSQ se utilizan para la reducción de ítems del instrumento inicial y una vez realizada la reducción, se deben obtener los valores de infit MNSQ y outfit MNSQ del cuestionario resultante de la reducción, con el fin de comprobar si los datos se ajustan al modelo de Rasch, de forma adecuada, pobre o excelente.

Los criterios para reducción de ítems están basados en unas pautas descritas por varios autores que se deben aplicar siguiendo el orden de prioridad siguiente:^{123,139,159,162}

1. Ítems con valores de infit MNSQ fuera del rango entre 0.70 y 1.30 logits.
2. Ítems con valores de oufit MNSQ fuera del rango entre 0.70 y 1.30 logits.
3. Ítems con alta proporción de datos perdidos (>50%).
4. Ítems con efecto techo: una alta proporción de respuestas en la categoría de respuesta más alta del ítem (>50%).
5. Ítems con una desviación estándar de la puntuación media considerablemente diferente de los otros ítems.
6. Ítems con coeficientes de asimetría y curtosis fuera del rango entre -2 y +2.

La reducción de ítems se realiza mediante un proceso iterativo de modo que se elimina un ítem cada vez, es decir, de uno en uno.¹⁵⁹ La secuencia de la eliminación es que el ítem con el mayor número de criterios que cumple, ordenados por prioridad, se elimina primero. Por lo tanto, cuando se elimina un ítem, el ajuste al modelo se vuelve a estimar hasta que se encuentre que ningún ítem debe ser eliminado. Se realiza de este modo puesto que se ha demostrado que este ajuste es relativo, de forma que la eliminación de ítems conduce a variaciones en el ajuste.

Este análisis de ajuste al modelo no es el único análisis que se utiliza para reducir ítems. Una vez que se realiza este proceso iterativo, se utilizan también el análisis de la dependencia local, para explorar si un ítem tiene dependencia de otros ítems, y el análisis del funcionamiento diferencial de ítems, para examinar si alguno de los

ítems funciona de manera diferente entre grupos. Si no cumplen estos requisitos, los ítems también son candidatos a ser eliminados.

C. Dependencia local

La dependencia local determina si la respuesta de cualquier ítem tiene un impacto directo en la respuesta de otro ítem.¹⁶⁶ La independencia local es un requisito del modelo, por lo que se debe realizar el análisis de la dependencia de los ítems a través del estadístico Yen-Q3, análisis que utiliza la matriz de correlación residual. Christensen et al,¹⁶⁶ han mostrado que no se puede considerar que un solo valor crítico indique dependencia local. Las simulaciones han demostrado que el valor crítico Q3 parece ser estable alrededor de un valor de 0.2 por encima de la correlación media. Por tanto, cualquier correlación residual mayor de 0.2 por encima de la correlación media indicaría dependencia local. Si dos ítems tienen dependencia local (cuando las correlaciones residuales de los ítems exceden la media de toda la correlación residual en 0.2), se debe conservar el ítem con un valor de outfit MNSQ más cercano a 1.¹⁶⁷

D. Análisis del funcionamiento diferencial de los ítems (DIF)

El análisis del funcionamiento diferencial de los ítems (Differential Item Functioning, DIF) se utiliza para determinar si un ítem mide las respuestas de un cuestionario de la misma manera para diferentes grupos de sujetos.¹¹⁷ Representa las diferencias en la dificultad de los ítems entre los grupos de encuestados y es importante evaluarlo ya que el DIF puede afectar el ajuste al modelo y puede perjudicar las medidas del cuestionario.¹⁴¹ Varios autores han mostrado que las

diferencias medias de las medidas de las personas entre los grupos comparados deben ser inferiores a 1.0 logit,¹²² de modo que cuando el valor sea mayor a 1.0 logits se debe considerar un DIF notable.¹¹⁶ Para que un ítem sea considerado con DIF, además de tener una diferencia de más de 1.0 logit, se debe determinar la significancia del contraste ($p < 0.05$) mediante el método de prueba t de Rasch-Welch.¹¹⁷ Si se encuentra DIF durante este análisis en alguno de los ítems, éstos son candidatos a ser eliminados.

E. Fiabilidad de la separación de las medidas de las personas y de los ítems

El análisis Rasch también proporciona lo que se conoce como fiabilidad (Reliability) o índice de separación (Separation index) tanto para las personas como para los ítems, análisis que muestra el rendimiento general del cuestionario.¹³⁹ Este valor de Reliability o Separation index indica cómo de consistente sería la medida de las personas o los ítems al medir el mismo constructo. En el caso de las personas, indica cómo de consistentemente se obtendrían estimaciones similares de la habilidad de las personas cuando se proporciona un conjunto de ítems.¹¹⁷ Y en el caso de los ítems, representa cómo sería la reproducibilidad de la jerarquía de los ítems si esos mismos ítems fueran respondidos por un grupo diferente de personas con una distribución similar de sus habilidades.¹¹⁷

La fiabilidad (Reliability) es una medida relativa con valores que oscilan entre 0 y 1 y el índice denominado Separation index es una medida continua cuyo valor se expresa en logits.

En el caso de las personas, el valor de Reliability que se considera aceptable es un resultado > 0.80 , y para Separation index > 2 logits. Esto significa que el

cuestionario es capaz de estratificar a la población en al menos tres grupos.¹⁶⁸ En el caso de un cuestionario de síntomas, significaría que el cuestionario es capaz de estratificar a los sujetos en al menos tres niveles de sintomatología. Si se obtiene un valor de Reliability o Separation index bajo para las personas, esto implica que el cuestionario necesitaría más ítems para poder estratificar adecuadamente a las personas.

En cuanto a los ítems, el valor de Reliability que se considera aceptable es un resultado superior a 0.90, que es equivalente a un Separation index >3 logits.^{139,169} Este valor indicaría que existe una correcta jerarquización de los ítems (un rango amplio de dificultades), pudiendo distinguir al menos cuatro niveles de estratificación de la dificultad de los ítems. Si se obtiene un valor de Reliability o Separation index bajo para los ítems, esto indicaría que la muestra no es lo suficientemente grande para mostrar la jerarquía de la dificultad de los ítems, por lo que se necesitaría aumentar el tamaño de la población.

F. Unidimensionalidad

La unidimensionalidad se explora en el análisis Rasch mediante el análisis de componentes principales (Principal Component Analysis, PCA) de los residuales, considerando residuales las diferencias entre los datos observados y la estimación del modelo.¹⁴¹ La medida se expresa en valores propios o autovalores (en inglés, eigenvalue), y la magnitud del primer contraste del residual, que es el indicador más importante, no debe tener un resultado por encima de 2 eigenvalue (autovalor), además de que la varianza explicada por las medidas debe ser mayor del 50%.¹⁴¹

G. Medida del alineamiento entre ítems y personas (targeting)

El análisis Rasch examina lo que se conoce como targeting mediante el mapa de personas-ítems.¹⁷⁰ El targeting es la diferencia entre la media de las habilidades de las personas y la media de las dificultades de los ítems. Cuanto más cerca esté la media de las habilidades de las personas a la media de la dificultad de los ítems, mejor será el targeting. Una diferencia de cero entre ambos valores indica un targeting perfecto de la escala y una diferencia de más de 1 logit indica un targeting deficiente.¹⁶²

H. Tabla de transformación de las puntuaciones

El análisis Rasch también permite transformar las puntuaciones ordinales del cuestionario a una escala de intervalo sin modificar la respuesta original del instrumento.¹⁴¹ Esta tabla de conversión es recomendable mostrarla, y se obtiene considerando el rango de puntuaciones ordinales de la escala, para después obtener las puntuaciones correspondientes a nivel de intervalo en logits y el rango de escala ordinal.¹⁴¹ La tabla de transformación puede ayudar a otros autores a utilizar el cuestionario para fines clínicos como poder comparar los cambios en la variable medida, ya que con una escala de intervalos iguales se pueden detectar variaciones en dicha medida. Esta conversión a puntuación de intervalo (medida en logits) facilita el uso de estadística paramétrica. Hay que tener en cuenta que para utilizar los valores de la tabla de transformación no deben faltar datos, por lo que todos los ítems deben ser completados por el sujeto evaluado. Para comparativa entre sujetos, la población debe ser similar a la utilizada en el estudio original.¹⁴¹

1.1.5.4. Exactitud diagnóstica mediante curvas ROC

Para conocer la exactitud diagnóstica de un cuestionario se debe realizar el análisis conocido como análisis de las curvas ROC (Receiver Operating Characteristic, ROC). Una curva ROC expresa la relación entre la sensibilidad (S) y la especificidad (Sp).⁴⁸ La sensibilidad en este caso es la capacidad del cuestionario para detectar los síntomas en sujetos con disfunciones y la especificidad se refiere a la capacidad del cuestionario en detectar la ausencia de síntomas en sujetos sin disfunción. Las curvas ROC representan para cada valor de corte, la tasa de verdaderos positivos (S) versus la tasa de falsos positivos (1 - Sp). De este modo, los valores de los ejes de las curvas ROC se sitúan entre 0 y 1. El punto de equilibrio entre la S y la Sp puede utilizar para decidir el punto de corte óptimo de la prueba, teniendo en cuenta que cuanto más cerca esté la curva al extremo superior izquierdo, mayor exactitud diagnóstica tendrá la prueba ya que aumentando la sensibilidad se pierde poca especificidad (en este caso de esta tesis doctoral la prueba sería el cuestionario que se está evaluando). De este modo, la exactitud diagnóstica de la prueba se describe como el área bajo la curva ROC, de modo que cuanto mayor sea el área, mejor será la prueba.⁴⁸ Un nivel de significación menor de 0.05 significará que el área bajo la curva ROC difiere del valor verdadero de 0.5.

1.1.5.5. Validez

Al evaluar un cuestionario, la validez indica hasta qué punto la herramienta mide el concepto que se pretende medir. La validez puede ser evaluada de varias formas, como la validez concurrente, validez de grupos conocidos o validez convergente y discriminante. Cabe resaltar que no existe una prueba única para analizar la validez, de modo que la validez de un cuestionario puede obtenerse mediante diferentes parámetros.

A. Validez concurrente

La validez concurrente es el análisis que se refiere al grado en que la puntuación de un cuestionario se correlaciona con la puntuación de medidas clínicas (en el caso de esta tesis doctoral serían los resultados de las pruebas visuales) mediante un análisis de correlación. Se considera una correlación adecuada cuando los valores varían entre 0.3 y 0.9.¹⁴⁰

B. Validez para grupos conocidos

La validez de grupos conocidos es la medida en que el instrumento puede discriminar entre grupos clínicamente diferentes que, teóricamente, deberían obtener resultados distintos al contestar el cuestionario.¹⁴⁰ Si un cuestionario es capaz de obtener diferentes resultados en grupos que clínicamente deberían serlo, esto indicaría una buena validez del mismo. Para ello, se deben comparar las puntuaciones del cuestionario obtenida por los grupos mediante pruebas de comparación para grupos independientes (T-test o U de Mann-Whitney).

C. Validez convergente y discriminante

La validez convergente implica que un instrumento se correlaciona con otro instrumento que mide un constructo similar.¹³⁹ En este caso, la correlación entre las puntuaciones de ambos instrumentos debe estar entre los valores de 0.3 y 0.9.^{139,140} La validez discriminante examina la medida en que el instrumento se correlaciona con un instrumento existente que mide un constructo diferente, y debe haber una correlación entre las puntuaciones menor de 0.3.^{139,140}

1.1.5.6. Repetibilidad

La repetibilidad es la consistencia del instrumento al realizar el mismo cuestionario en diferentes administraciones.¹³⁹ Se determina cuando se quiere valorar un análisis test-retest, que indica el grado en que los resultados son repetibles cuando son realizados por el mismo observador.¹³⁹ Para ello, se recomienda que se realice la segunda medida con un tiempo entre ellas de al menos una semana después de la primera medida.¹⁷¹ El método más utilizado para realizar un test-retest es el índice estadístico conocido como coeficiente de correlación intraclase (Intraclass Correlation Coefficient, ICC), que se define como la proporción de variabilidad total debida a la variabilidad de los pacientes.¹⁷² El valor del ICC oscila entre 0 y 1, y se debe calcular con sus correspondientes intervalos de confianza al 95%. La literatura científica recomienda un valor de ICC igual o mayor de 0.80 para considerar que el instrumento tiene una excelente repetibilidad.^{139,140,172}

Todas estas medidas para la validación de un cuestionario, como se ha comentado anteriormente, son recomendaciones. Estas recomendaciones añaden calidad a la elaboración y validación de un cuestionario, por lo que es importante en dicho proceso realizar todas las que se puedan, además de elegir los tests estadísticos más recomendados y utilizados en la actualidad.

1.2. OBJETIVOS

El objetivo general de esta investigación es desarrollar y validar un cuestionario sobre la sintomatología asociada a las disfunciones visuales, que pueda ser utilizado en una población clínica. De forma más precisa, los objetivos específicos son:

- a. Identificar los síntomas que están asociados a cualquier disfunción visual, ya sea refractiva, acomodativa o binocular, y cuáles deberían incluirse en un cuestionario sobre síntomas visuales.
- b. Desarrollar un cuestionario de síntomas asociados a las disfunciones visuales, utilizando una población clínica.
- c. Valorar las propiedades psicométricas de este cuestionario, incluyendo un análisis de su exactitud diagnóstica, validez y repetibilidad en una población clínica.

1.3. HIPÓTESIS

Se plantean tres hipótesis de trabajo en esta investigación:

1. Las disfunciones visuales presentan una serie de síntomas que pueden ser comunes entre los diferentes tipos de anomalías. Estos síntomas se encuentran en la práctica clínica habitual, y para su conocimiento, puede ser útil la opinión de expertos optometristas del ámbito clínico.
2. Conociendo los síntomas asociados a las disfunciones visuales se puede diseñar y construir un cuestionario sobre sintomatología visual que pueda utilizarse en la población clínica.
3. La puntuación de este cuestionario es capaz de detectar la presencia de síntomas en sujetos con disfunciones visuales.

Las hipótesis de trabajo se relacionan con el diseño, construcción y uso clínico de un cuestionario para las disfunciones visuales.

La primera de las hipótesis sugiere que la sintomatología visual producida por las disfunciones visuales no tiene por qué ser exclusiva de una anomalía en concreto, pudiendo ser compartida por varias disfunciones.

La segunda de ellas hace referencia a que es posible construir un instrumento con el que poder valorar la sintomatología de estas disfunciones, para lo que es necesario previamente establecer qué síntomas pueden utilizarse conjuntamente en un cuestionario específico para poder posteriormente utilizarlo en una población diana, que es la población clínica.

La tercera hipótesis considera que un cuestionario sobre sintomatología permite obtener una puntuación con la que se puede asociar un determinado valor a la existencia de síntomas provocados por cualquier tipo de disfunción visual. De este modo, dicho cuestionario podría utilizarse como ayuda en el diagnóstico y tratamiento de la sintomatología asociada a estas anomalías, además de poder utilizarlo en estudios de investigación.

1.4. JUSTIFICACIÓN

Como se ha podido observar, no existe un cuestionario de síntomas diseñado con preguntas que puedan relacionarse con cualquier tipo de disfunción visual. Debido a ello, puede resultar interesante y útil disponer de un cuestionario que permita la detección de cualquier sintomatología visual asociada a las disfunciones visuales. Este cuestionario podría ser utilizado por los profesionales para concretar la sintomatología del paciente tanto en el diagnóstico de estas anomalías como en la monitorización de su tratamiento.

En la literatura científica se han utilizado en algunos estudios los cuestionarios de Conlon¹³⁰ y CISS^{55,155,156} con el fin de recoger la sintomatología que producen las disfunciones visuales.^{30,62,173-180} Como se ha comentado anteriormente, el cuestionario de Conlon¹³⁰ se ha desarrollado para medir el disconfort visual, pero el disconfort visual no es exactamente lo mismo que los síntomas que pueden aparecer en las disfunciones visuales. De la misma manera, el cuestionario CISS^{55,155,156} es un cuestionario de síntomas, pero ha sido validado únicamente para la IC, por lo que utilizarlo en otras condiciones no es adecuado ya que está bien establecido en el ámbito epidemiológico,

que un cuestionario solo debería ser utilizado para la condición para la cual fue desarrollado y validado.^{181,182} De hecho diversos autores han encontrado dificultades a la hora de utilizar dicho cuestionario. De este modo, se ha observado que sujetos con visión binocular normal obtienen puntuaciones altas con el instrumento CISS.¹⁸³ Otros autores han encontrado que estos síntomas asociados con la IC también son comunes en sujetos sin esta anomalía binocular.¹⁸⁴ Asimismo se ha encontrado que el CISS sobreestima los síntomas en visión próxima por lo que no debería utilizarse como una herramienta de cribado para la IC.¹⁸⁵ Otro estudio sugiere que el cuestionario CISS no es unidimensional,¹⁸⁶ e incluso algunas investigaciones han encontrado una baja especificidad de la herramienta.^{59,62,179} Además, hay que tener en cuenta que la herramienta CISS no se ha validado mediante un análisis psicométrico, sino tan solo mediante un análisis de las curvas ROC, por lo que este hecho debería tenerse en cuenta a la hora de su utilización.

Como se ha comentado anteriormente, al no existir un cuestionario específico de sintomatología visual asociada a las disfunciones visuales, el objetivo de esta tesis doctoral ha sido desarrollar y validar un cuestionario con el que recoger los síntomas asociados a cualquier tipo de disfunción visual. Para poder alcanzar los objetivos específicos, se han realizado tres artículos científicos, cada uno correspondiente con cada uno de los objetivos específicos, en los cuales se han tenido en cuenta las recomendaciones existentes en la literatura científica para el desarrollo y validación de cuestionarios.¹³⁹⁻¹⁴¹ En el primer artículo, se realiza la identificación de los ítems a incluir en un cuestionario, en el segundo la selección de los ítems incluidos en el cuestionario mediante una reducción psicométrica y un análisis exploratorio sobre las características psicométricas del cuestionario, y en el tercero, se realiza un análisis confirmatorio de las propiedades del cuestionario mediante análisis psicométrico, además de evaluar la exactitud diagnóstica, la validez y la repetibilidad.

1.5. RESUMEN GLOBAL DE LOS RESULTADOS OBTENIDOS Y DISCUSIÓN

Con el fin de alcanzar los objetivos de esta tesis doctoral, se desarrollaron tres investigaciones, cada una correspondiente a un objetivo específico y que representan cada uno de los subapartados de esta sección.

Esta investigación fue aprobada por el Comité de Ética de la Universidad de Alicante (Anexo 1). El estudio siguió los principios de la Declaración de Helsinki, y todos los métodos de esta investigación se llevaron a cabo de acuerdo con las directrices y normativas de esta declaración.

En la primera de las investigaciones, titulada “Metodología Delphi para la detección de sintomatología asociada a las disfunciones visuales”, se realizó una metodología Delphi para analizar qué opinan los profesionales de la visión sobre qué síntomas se asocian a las disfunciones refractivas, acomodativas o binoculares, y cuáles de estos síntomas deberían utilizarse en un cuestionario sobre sintomatología visual.

En la segunda, titulada “Análisis Rasch para el desarrollo y reducción del Symptom Questionnaire for Visual Dysfunctions (SQVD)”, y basándose en los resultados obtenidos en la primera sección, se desarrolló un cuestionario para detectar la presencia y frecuencia de la sintomatología visual asociada a cualquier disfunción visual. En esta investigación se realizó un análisis Rasch para hacer una reducción psicométrica del número de ítems y posteriormente su análisis psicométrico.

Partiendo de los resultados de esta segunda investigación, se realizó una tercera investigación titulada “Validación del Symptom Questionnaire for Visual Dysfunctions (SQVD), un cuestionario para evaluar los síntomas de cualquier tipo de disfunción visual”, en la cual se validó psicométricamente el cuestionario SQVD en una muestra clínica de pacientes con cualquier tipo de disfunción visual, analizando también su exactitud diagnóstica, validez y repetibilidad.

1.5.1. Metodología Delphi para la detección de sintomatología asociada a las disfunciones visuales

Para iniciar el proceso de elaboración del cuestionario de sintomatología asociada a las disfunciones visuales, se decidió realizar una metodología Delphi con la que conocer la opinión de expertos sobre este tema. La aplicación del método Delphi presenta ciertas ventajas sobre otras técnicas de consenso, como permitir que todos los expertos tengan una participación e influencia iguales, así como poder reunir a diferentes expertos de la geografía nacional sin tener que reunirlos presencialmente. Debido a la falta de disponibilidad de un cuestionario de síntomas para evaluar cualquier disfunción visual, esta investigación tuvo como objetivo utilizar un enfoque cualitativo para localizar los síntomas asociados a las anomalías visuales y determinar cuáles podrían incluirse en un cuestionario sobre sintomatología visual. La evidencia científica¹¹² ha mostrado que los cuestionarios existentes no se han desarrollado teniendo en cuenta la experiencia de los profesionales del ámbito clínico optométrico. Por todo ello, utilizar un método Delphi para saber qué síntomas son considerados por los clínicos para desarrollar un cuestionario sobre sintomatología asociada con las disfunciones visuales puede ser una ayuda importante basada en la evidencia científica.

La metodología Delphi desarrollada estableció un grupo coordinador y un grupo de expertos. El grupo de coordinación estuvo integrado por Ángel García Muñoz, Pilar Cacho Martínez y Mario Cantó Cerdán, todos ellos optometristas experimentados y expertos en salud pública, que fueron los responsables de la coordinación general de la logística del estudio. El rol del grupo coordinador abarcó funciones como modificar el protocolo del estudio cuando fuera necesario, reclutar expertos, preparar los cuestionarios de cada ronda, analizar sus respuestas e interpretar los resultados del estudio.

El grupo de expertos fueron los encargados de dar su opinión en cada ronda. Inicialmente se invitó a 17 optometristas españoles de distintas áreas geográficas a participar en el estudio, todos ellos profesionales de reconocido prestigio con más de 15 años de experiencia tanto en el campo de la evaluación como del tratamiento de diferentes anomalías visuales. Tres de ellos rechazaron la invitación, por lo que el panel final de expertos estuvo formado por 14 miembros (Anexo 2). Los expertos fueron seleccionados atendiendo a sus conocimientos y a su experiencia, que son las principales características recomendadas¹⁵⁶ para considerar un experto en una técnica Delphi. De todos ellos, 5 optometristas eran profesionales que desarrollaban su práctica clínica en clínicas oftalmológicas y los otros 9 en establecimientos sanitarios de óptica, por lo que esto permitió tener diferentes puntos de vista. Todos los expertos atendían a pacientes de todos los grupos de edad y de diferentes áreas de la salud visual en sus prácticas clínicas habituales. Fueron contactados individualmente por correo electrónico para preservar el anonimato. Todos los expertos dieron su aceptación por consentimiento informado para participar en la investigación.

Para recopilar la información obtenida en el panel de expertos, se diseñó una base de datos específica. Al ser expertos de diferentes áreas geográficas, este método permitió obtener diferentes ideas de varios optometristas de toda España y a su vez, debido al anonimato, se evitaron posibles influencias negativas de los miembros dominantes del grupo o la inhibición de algún participante.^{153,154}

El grupo coordinador fue el encargado de formular la pregunta de investigación al panel de expertos de este método Delphi, el cual se llevó a cabo en tres rondas. En cada una de las rondas había un intercambio de información entre el grupo coordinador y cada experto individualmente. Una vez enviada la pregunta de cada ronda al grupo de expertos, el grupo coordinador recogía y analizaba las opiniones y contestaciones de los expertos. En base a la interpretación y análisis de estas respuestas, el grupo

coordinador desarrollaba nuevas preguntas que se plasmaban en un cuestionario que de nuevo se enviaba a los expertos en la siguiente ronda. Además, el grupo coordinador enviaba un informe con el resumen de los resultados obtenidos al final de cada ronda y así, antes de iniciar la siguiente ronda, los expertos conocían las opiniones de todos ellos para que pudieran reafirmar sus respuestas. Este método permitió a los expertos la posibilidad de reflexionar o reconsiderar su opinión a través de sus propios puntos de vista o de los otros expertos. Cada ronda duró dos semanas, y si los participantes no respondían en este momento o no completaban el cuestionario, se les enviaba un correo electrónico recordatorio, evitando así sesgos en el estudio.

1.5.1.1. Primera ronda de la metodología Delphi

En la primera ronda, se recopiló una lista de 34 categorías de síntomas (Tabla 4) proporcionadas por la revisión sistemática de García-Muñoz et al¹¹² sobre síntomas y anomalías visuales. En esta ronda, se les preguntó a los expertos qué síntomas de estos 34 consideraban que podían estar relacionados con las disfunciones visuales, pidiéndoles que los clasificaran según el tipo de anomalía (refractiva, acomodativa o binocular). Además, también se les preguntó sobre la distancia de trabajo (visión lejana o visión próxima) con la que se podría relacionar cada uno de estos síntomas. Para considerar todas estas relaciones, se tuvo en cuenta que las respuestas de los expertos fueran designadas por el 50% o más de los participantes, es decir, cuando 7 o más expertos consideraban que un determinado síntoma se relacionaba con el tipo de anomalía o la distancia de trabajo, el síntoma se consideraba que tenía una relación importante.

Teniendo en cuenta estas consideraciones, se obtuvo que, según los expertos, 28 de los 34 síntomas iniciales (82.4%) estaban relacionados con las disfunciones

refractivas (Figura 1). Los síntomas que los expertos asociaron más a las anomalías refractivas fueron *astenopía, dolor de cabeza, visión borrosa, dificultad para realizar las tareas escolares, disconfort visual, fatiga visual y problemas de lectura*. Para las disfunciones acomodativas, 23 de los 34 síntomas (67.6%) los asociaron con estas anomalías acomodativas (Figura 2). Los síntomas *dificultad para cambios de enfoque astenopía, dolor de cabeza, falta de concentración, problemas de lectura y visión borrosa* son los que los seis síntomas que los expertos más relacionaron con las disfunciones acomodativas. Cabe resaltar que cuatro de estos seis síntomas (*astenopía, dolor de cabeza, problemas de lectura y visión borrosa*) coinciden con los de las disfunciones refractivas. Para las disfunciones binoculares, los expertos señalaron que 28 de los 34 síntomas (82.4%) estarían relacionados con disfunciones binoculares (Figura 3), siendo los síntomas más relacionados *diplopía, notar que un ojo se desvía y calcular mal las distancias*. Además, los expertos también asociaron a las disfunciones refractivas muchos de los síntomas relacionados con las disfunciones acomodativas, por lo que se puede observar que consideraban que los síntomas para ambos casos podían ser los mismos. Este resultado es similar al obtenido en el estudio de Cacho-Martínez et al,¹¹¹ donde los autores encontraron que los errores refractivos no corregidos pueden enmascarar los síntomas de las disfunciones acomodativas y binoculares, fundamentalmente los de las disfunciones acomodativas.

Los resultados de la primera ronda también mostraron que los expertos relacionaron con la visión lejana a 11 de los 34 síntomas (Figura 4), siendo los síntomas *calcular mal las distancias, diplopía y visión borrosa* los más destacados por los expertos. Sin embargo, la mayoría de los síntomas (26 de los 34 síntomas) los asociaron a la visión próxima (Figura 5), destacando los síntomas de *diplopía, disconfort visual, fatiga visual, astenopía, dificultad para realizar tareas escolares, dolor de cabeza, falta de concentración, movimiento o parpadeo del texto en cerca, pesadez ocular, problemas*

de lectura y visión borrosa. De estos síntomas, hay dos que son compartidos con la visión lejana (diplopía y visión borrosa).

En general, también se encontró que la mayoría de los síntomas relacionados con la visión próxima coincidían con los síntomas más relacionados con cada una de las disfunciones visuales en particular. Estos resultados son similares a los obtenidos en la revisión sistemática de García-Muñoz et al¹¹² en la que los autores demostraron que la mayoría de los síntomas reflejados por la literatura científica estaban asociados a la visión próxima.

En esta primera ronda los expertos sugirieron añadir 10 síntomas a los 34 síntomas iniciales. Estos síntomas se consideraron cuando al menos uno de los expertos recomendó su inclusión. Los 10 síntomas adicionales que se incluyeron propuestos por seis de los expertos fueron: *bruma / telaraña; dificultad de fijación; dificultad para terminar las actividades propuestas a tiempo; dolor de cuello y/o espalda; dolor en la zona de las cejas; entornar los ojos para ver mejor; incapacidad de mantener actividades exigentes de esfuerzo visual de manera continua; irritabilidad frente a actividades exigentes de esfuerzo visual; olvidadizo y torpeza / se choca con objetos.*

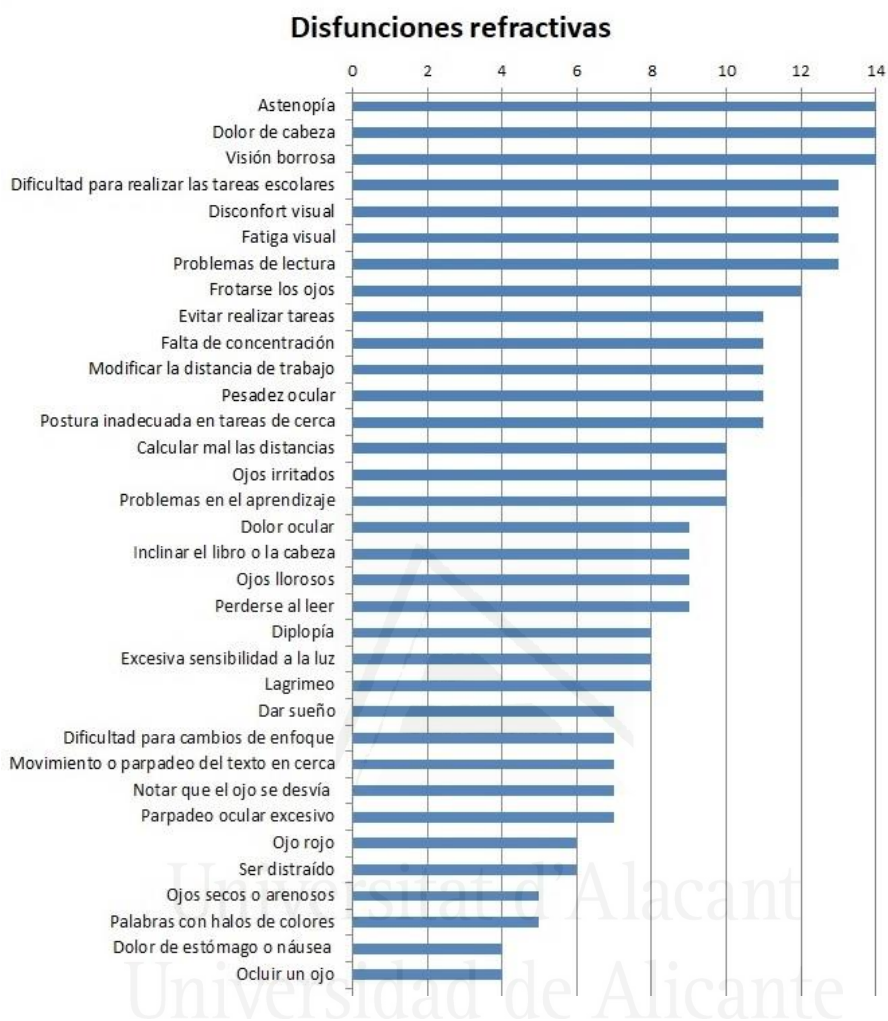


Figura 1. Resultados de la primera ronda para las disfunciones refractivas. La barra de frecuencia indica el número de expertos que asociaban cada uno de los 34 síntomas iniciales a las disfunciones refractivas.

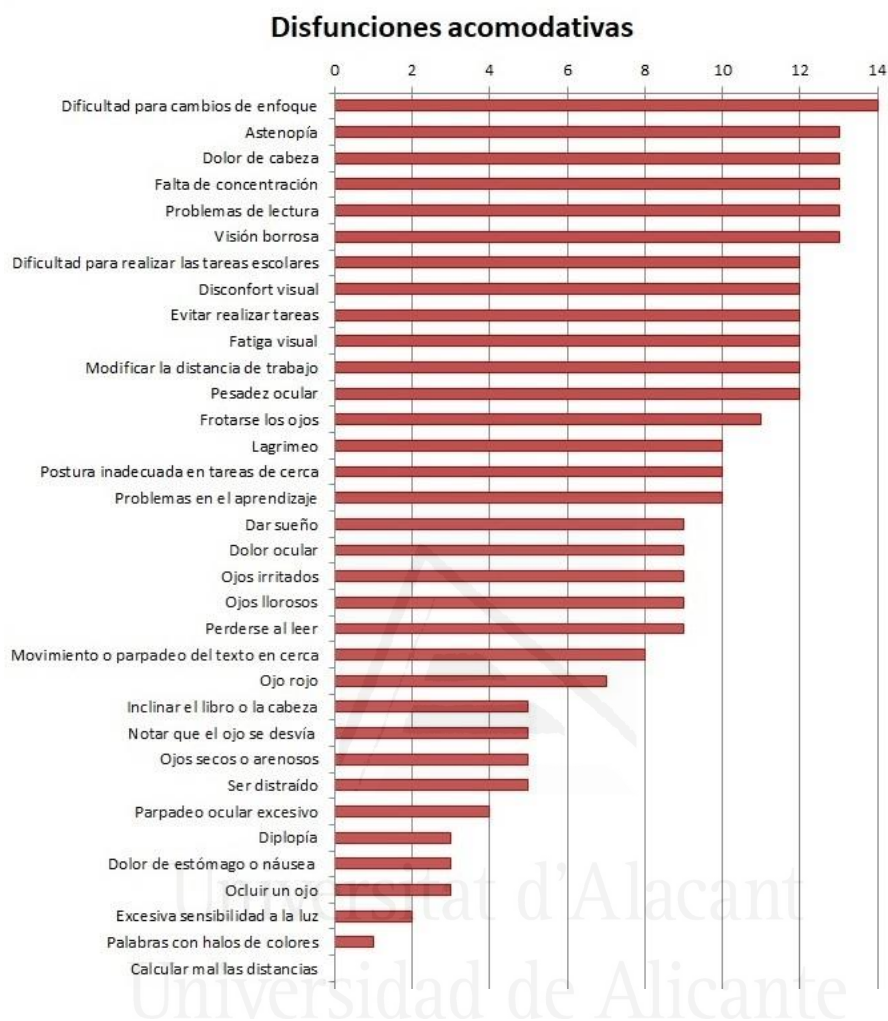


Figura 2. Resultados de la primera ronda para las disfunciones acomodativas. La barra de frecuencia indica el número de expertos que asociaban cada uno de los 34 síntomas iniciales a las disfunciones acomodativas.

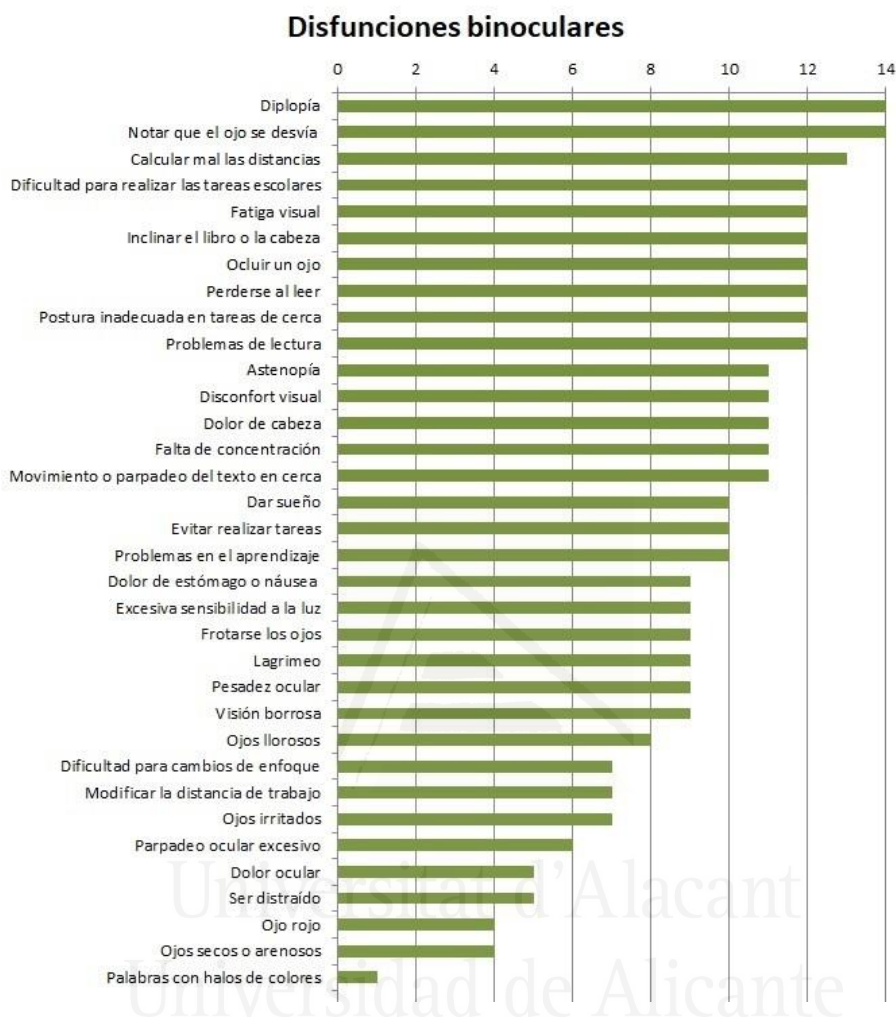


Figura 3. Resultados de la primera ronda para las disfunciones binoculares. La barra de frecuencia indica el número de expertos que asociaban cada uno de los 34 síntomas iniciales a las disfunciones binoculares.

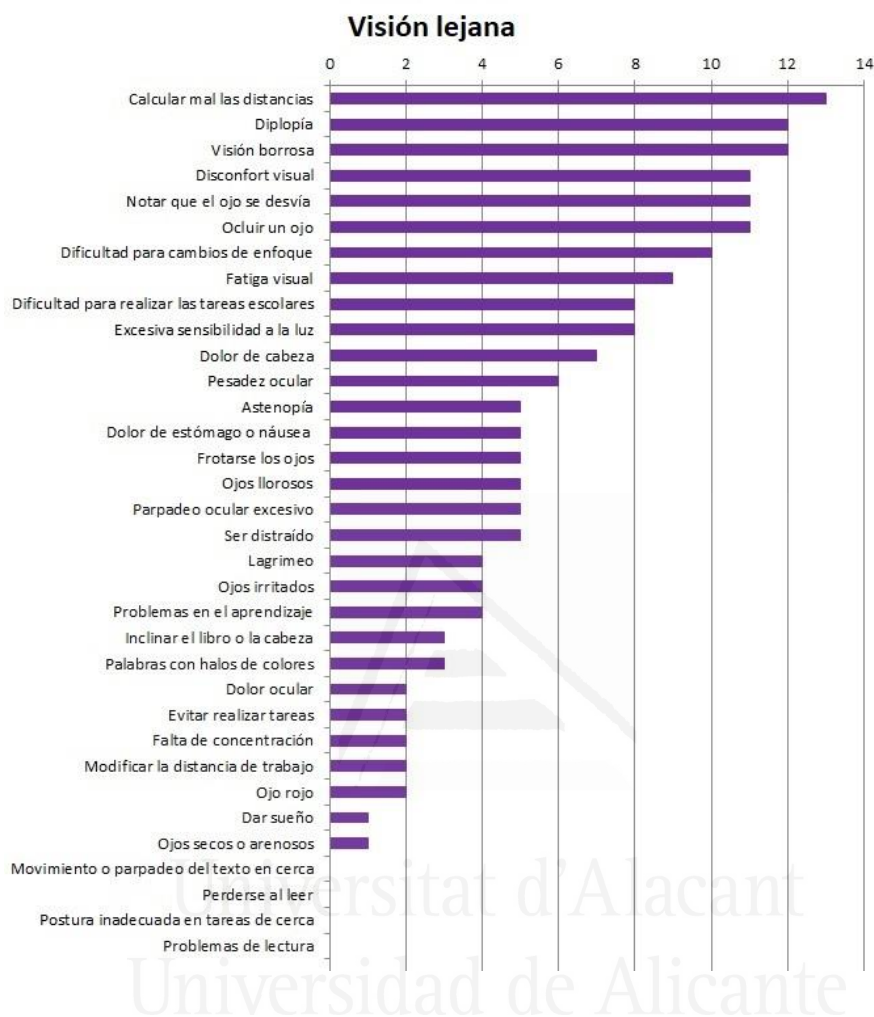


Figura 4. Resultados de la primera ronda para la visión lejana. La barra de frecuencia indica el número de expertos que asociaban cada uno de los 34 síntomas iniciales a la visión lejana.

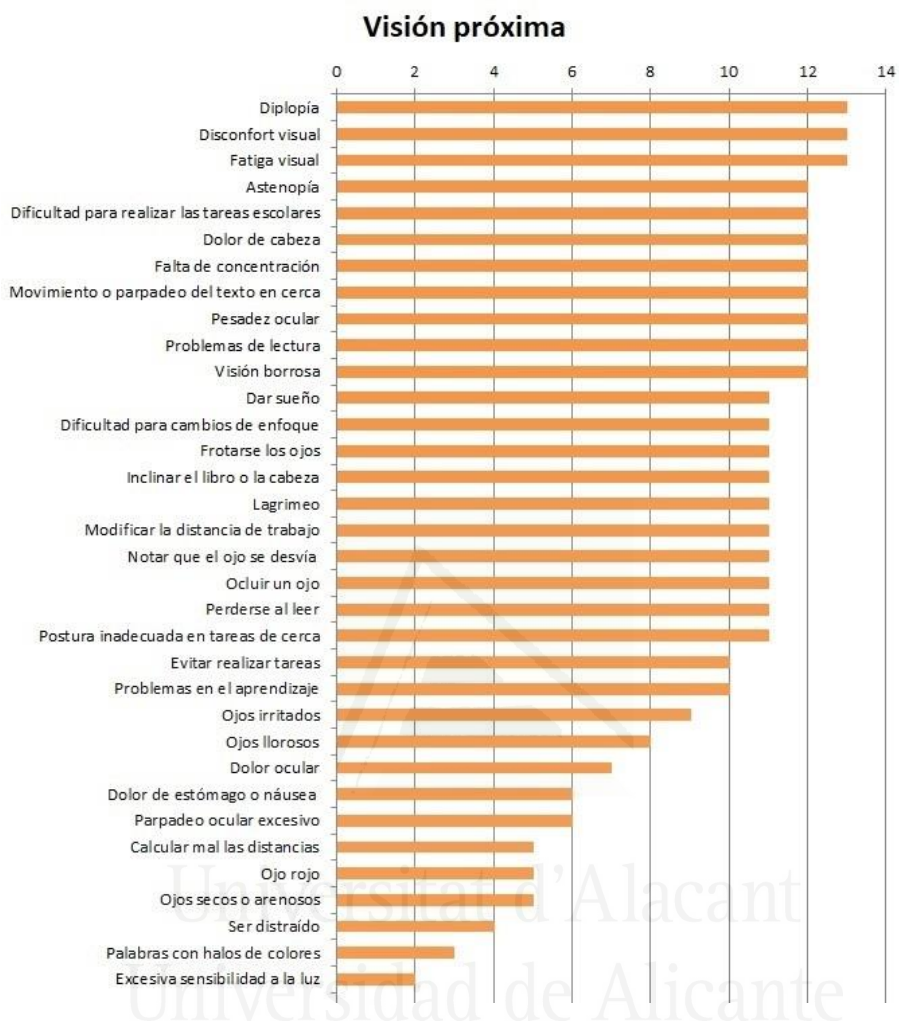


Figura 5. Resultados de la primera ronda para la visión próxima. La barra de frecuencia indica el número de expertos que asociaban cada uno de los 34 síntomas iniciales a la visión próxima.

1.5.1.2. Segunda ronda de la metodología Delphi

En la segunda ronda se les pidió a los expertos que cuantificaran la relación entre las nuevas 44 categorías de síntomas (pues se habían añadido diez síntomas a los iniciales propuestos) y cada una de las disfunciones. Los expertos debían calificar la relación en una escala de 0 a 10, siendo 0 la puntuación menos relacionada y 10 la más relacionada. Una vez recibidas todas las respuestas, se sintetizaron y analizaron (mediante la media y la desviación estándar), resultados mostrados en la Tabla 5.

Como se observa, para las disfunciones refractivas, la *visión borrosa* obtuvo la puntuación más alta (9.64 ± 0.61), aunque a su vez también recibió una puntuación alta en las disfunciones acomodativas (7.79 ± 2.14), lo que mostraba que los expertos no consideraban a este síntoma exclusivo de las anomalías refractivas,¹⁸⁷ pudiendo relacionarse con una anomalía refractiva o acomodativa.^{51,111,188,189} Esto concuerda con el hecho de que el síntoma de *visión borrosa* aparece en la mayoría de los cuestionarios visuales mostrados por la literatura científica.^{55,138,152-161} También se observó que, para las disfunciones refractivas, los síntomas de *astenopía* y *entornar los ojos para ver mejor* fueron los otros síntomas que obtuvieron las puntuaciones más altas (8.93 ± 1.33 y 9.36 ± 0.97 respectivamente), siendo este último un síntoma que los expertos otorgaron una puntuación muy alta en esta disfunción, pero no para el resto, lo que sugeriría que este síntoma podría ser específico para las disfunciones refractivas. Curiosamente, este síntoma de *entornar los ojos para ver mejor* fue uno de los diez síntomas que agregaron los expertos en la primera ronda. Sin embargo, la evidencia científica ha demostrado que solo aparece en dos cuestionarios, el cuestionario de Conlon¹³⁰ y el descrito pero no validado por Adler et al.¹³³

	Disfunciones refractivas		Disfunciones acomodativas		Disfunciones binoculares	
	Media	DE	Media	DE	Media	DE
Astenopía	8.93	1.33	9.14	1.36	8.71	1.71
Bruma / Telaraña	7.29	1.67	5.14	2.82	2.21	2.86
Calcular mal las distancias	5.57	2.61	3.14	2.64	8.79	1.82
Dar sueño	5.50	3.35	6.14	3.48	5.43	3.22
Dificultad de fijación	5.57	3.22	7.00	1.85	8.07	1.33
Dificultad para cambios de enfoque	6.64	2.69	9.43	0.73	5.64	3.15
Dificultad para realizar las tareas escolares	8.14	1.41	8.21	1.70	8.21	1.61
Dificultad para terminar las actividades propuestas a tiempo	6.93	1.71	7.64	2.02	7.64	1.95
Diplopía	4.00	2.88	3.50	3.68	9.14	1.06
Disconfort visual	8.14	1.41	8.29	1.79	8.43	1.40
Dolor de cabeza	7.92	3.22	7.79	2.04	8.21	1.78
Dolor de cuello y/o espalda	3.64	3.08	3.50	2.87	5.57	2.87
Dolor de estómago o náusea	4.29	2.37	4.57	3.20	6.86	2.87
Dolor en la zona de las cejas	6.00	2.20	5.57	3.06	4.86	3.00
Dolor ocular	5.36	2.35	5.93	2.74	4.93	2.89
Entornar los ojos para ver mejor	9.36	0.97	6.36	3.24	3.21	3.07
Evitar realizar tareas	7.36	2.06	7.79	1.82	7.57	2.06
Excesiva sensibilidad a la luz	5.50	3.31	2.93	2.63	7.36	3.33
Falta de concentración	7.29	2.19	7.79	1.66	8.00	1.56
Fatiga visual	8.14	1.77	8.29	2.58	7.93	2.49
Frotarse los ojos	7.36	1.67	7.21	2.57	7.07	2.68
Incapacidad de mantener actividades exigentes de esfuerzo visual de manera continua	7.79	1.47	8.43	1.35	8.50	1.35
Inclinar el libro o la cabeza	5.00	2.27	4.00	2.45	8.36	1.76
Irritabilidad con actividades de esfuerzo visual	6.36	2.50	7.36	1.87	7.57	2.66
Lagrimo	6.14	2.36	6.29	2.43	6.36	2.64
Modificar la distancia de trabajo	8.21	1.37	8.07	2.52	6.93	2.52
Movimiento o parpadeo del texto en cerca	6.71	2.05	7.29	2.52	6.64	3.04
Notar que el ojo se desvía	4.00	3.40	3.50	2.99	9.57	0.73
Ocluir un ojo	4.29	3.30	3.71	2.68	9.71	0.59
Ojo rojo	5.36	2.22	5.14	2.67	4.00	2.88
Ojos irritados	5.71	2.68	5.43	2.90	4.93	3.39
Ojos llorosos	5.93	2.43	5.64	2.72	5.71	3.49
Ojos secos o arenosos	4.21	2.57	3.79	2.83	2.86	2.82
Olvidadizo	1.43	1.72	1.79	2.34	2.21	2.93
Palabras con halos de colores	4.29	3.30	3.21	3.41	3.07	3.24
Parpadeo ocular excesivo	5.71	2.08	5.79	2.24	6.71	2.91
Perderser al leer	5.50	2.97	6.07	2.68	8.21	1.82
Pesadez ocular	6.64	2.19	7.07	2.52	6.00	2.88
Postura inadecuada en tareas de cerca	6.50	2.56	7.36	2.35	8.21	2.04
Problemas de lectura	6.93	2.52	7.93	2.09	8.07	2.09
Problemas en el aprendizaje	5.79	3.71	6.21	3.78	6.36	3.64
Ser distraído	4.93	3.13	4.57	3.92	4.86	3.68
Torpeza; se choca con objetos	5.64	2.89	2.50	2.90	7.64	2.87
Visión borrosa	9.64	0.61	7.79	2.14	4.93	2.55

Tabla 5. Resultados de la segunda ronda, donde se indican la media y desviación estándar (DE) de las puntuaciones otorgadas por el panel de expertos, relacionando cada uno de los 44 síntomas con las disfunciones visuales (refractivas, acomodativas y binoculares). La escala de puntuación es de 0 a 10 puntos.

Para las anomalías acomodativas, el síntoma de *dificultad para cambios de enfoque* fue el síntoma al que los expertos le dieron la puntuación más alta (9.43 ± 0.73). Según la opinión de los expertos, este síntoma parece ser específico para estas anomalías, a pesar de que solo aparece en el cuestionario utilizado por Sterner et al.¹³⁶ El otro síntoma que también obtuvo una media alta, *astenopía* (9.14 ± 1.36), también recibió por parte de los expertos una puntuación alta para las demás disfunciones, mostrando que los expertos le daban importancia para todas las anomalías visuales, aunque en ningún caso el síntoma de astenopía aparece en ninguno de los cuestionarios existentes.^{34,128-138} Esta situación es lógica, pues los expertos trataron a la astenopía como una categoría y no como un síntoma concreto. De hecho, la astenopía es un conjunto de síntomas y, por tanto, si apareciera descrito de ese modo los pacientes no entenderían dicho término.

Para las disfunciones binoculares, los síntomas con las puntuaciones medias más altas fueron *ocluir un ojo* (9.71 ± 0.59), *notar que un ojo se desvía* (9.57 ± 0.73) y *diplopía* (9.14 ± 1.06). De estos tres, la *diplopía* aparece en ocho¹²⁸⁻¹³⁵ de los 11 cuestionarios relacionados con algún tipo de sintomatología visual.^{34,128-138} Sin embargo, el síntoma de *ocluir un ojo* solo aparece en el 19-Item College of Optometrist in Vision Development Quality of Life (COVD-QOL)¹³⁷ y el síntoma *notar que un ojo se desvía* no aparece en ningún cuestionario.

Haciendo un análisis general de esta segunda ronda, los síntomas *dolor de cuello y/o espalda*, *dolor de estómago o náusea*, *ojos secos o arenosos*, *olvidadizo*, *palabras con halos de colores* y *ser distraído* obtuvieron bajas puntuaciones medias para todas las anomalías. Este hallazgo puede deberse a la poca importancia que los expertos les dieron a esos síntomas. Además, los resultados esta segunda ronda mostraron que para las anomalías binoculares hubo 15 síntomas (un 34% del total) con puntuaciones medias superiores a 8. Este resultado es superior al encontrado para las

disfunciones refractivas y acomodativas, que cada una de ellas tenía 7 síntomas (un 15.91 % del total) con puntuaciones superiores a 8. Esto muestra que los expertos dieron más importancia a los síntomas asociados a las disfunciones binoculares que a las refractivas y acomodativas, resultado que está en concordancia con el encontrado por Cacho-Martínez et al,¹¹¹ donde, en una población clínica, se observó una mayor asociación entre tener síntomas y presentar una disfunción binocular. A su vez, los resultados obtenidos en esta investigación son similares a los hallados en la revisión sistemática de García-Muñoz et al,¹¹² donde la mayoría de los autores de los artículos científicos relacionaban los síntomas visuales con las disfunciones binoculares. En cualquier caso, según se ha podido observar, el mismo síntoma puede estar relacionado con varias disfunciones. Y aunque parece que los síntomas pueden estar más asociados con las disfunciones binoculares, esto no implica que sean específicos de estas anomalías. El hecho de considerar que no existan síntomas específicos para cada disfunción visual sugiere que el desarrollo de un cuestionario de sintomatología no debería enfocarse solamente a una disfunción en particular, sino que puede relacionarse con las diferentes disfunciones existentes.

1.5.1.3. Tercera ronda de la metodología Delphi

En la tercera ronda los expertos tuvieron que indicar si cada uno de los 44 síntomas debían aparecer en un cuestionario de síntomas visuales. Para ello, se utilizó una encuesta con una escala Likert de 5 respuestas, donde los participantes debían responder su nivel de acuerdo para incluir cada síntoma en un cuestionario de sintomatología visual (*muy en desacuerdo / en desacuerdo / ni de acuerdo ni en desacuerdo / de acuerdo / totalmente de acuerdo*). Los expertos solo tenían que señalar una de las respuestas específicas para cada síntoma, de modo que cuando se

recibieron todas las respuestas, se sintetizaron y analizaron (mediante la mediana y el rango intercuartílico expuesto a través de un diagrama de caja y bigotes). El grupo coordinador asumió el consenso para los síntomas que debían aparecer en un cuestionario cuando la mediana estaba en el nivel de *totalmente de acuerdo*, además de tener todas las puntuaciones entre *ni de acuerdo ni en desacuerdo* y *totalmente de acuerdo*. Utilizando este criterio, se puede asegurar que ningún experto estuvo en desacuerdo con la inclusión de un síntoma en el cuestionario. De forma análoga, se asumió el consenso para los síntomas que no debían aparecer en un cuestionario cuando la mediana se encontrara entre en el nivel de *muy en desacuerdo*, *en desacuerdo* y *ni de acuerdo ni en desacuerdo*, independientemente de que las puntuaciones pudieran situarse en niveles superiores.

Los resultados mostraron que los expertos consideraron que 15 síntomas sí deberían aparecer en un cuestionario, 5 síntomas no deberían estar y no hubo consenso para los 24 síntomas restantes. Estos resultados se pueden ver en la Figura 6 que muestra un diagrama de cajas. El ancho de cada caja muestra el rango intercuartílico (RIC). La línea naranja de cada caja marca la mediana de las respuestas (que es el percentil 50). El extremo izquierdo de cada caja muestra el cuartil inferior (percentil 25) y el extremo derecho representa el cuartil superior (percentil 75). El bigote izquierdo indica el valor más bajo obtenido (un valor más bajo dentro de 1.5 RIC) y el bigote derecho muestra el valor más alto obtenido (que es el valor más alto dentro de 1.5 RIC). El asterisco representa un valor atípico situado entre 1.5 y 3 RIC, y el círculo muestra un valor atípico superior a 3 RIC.

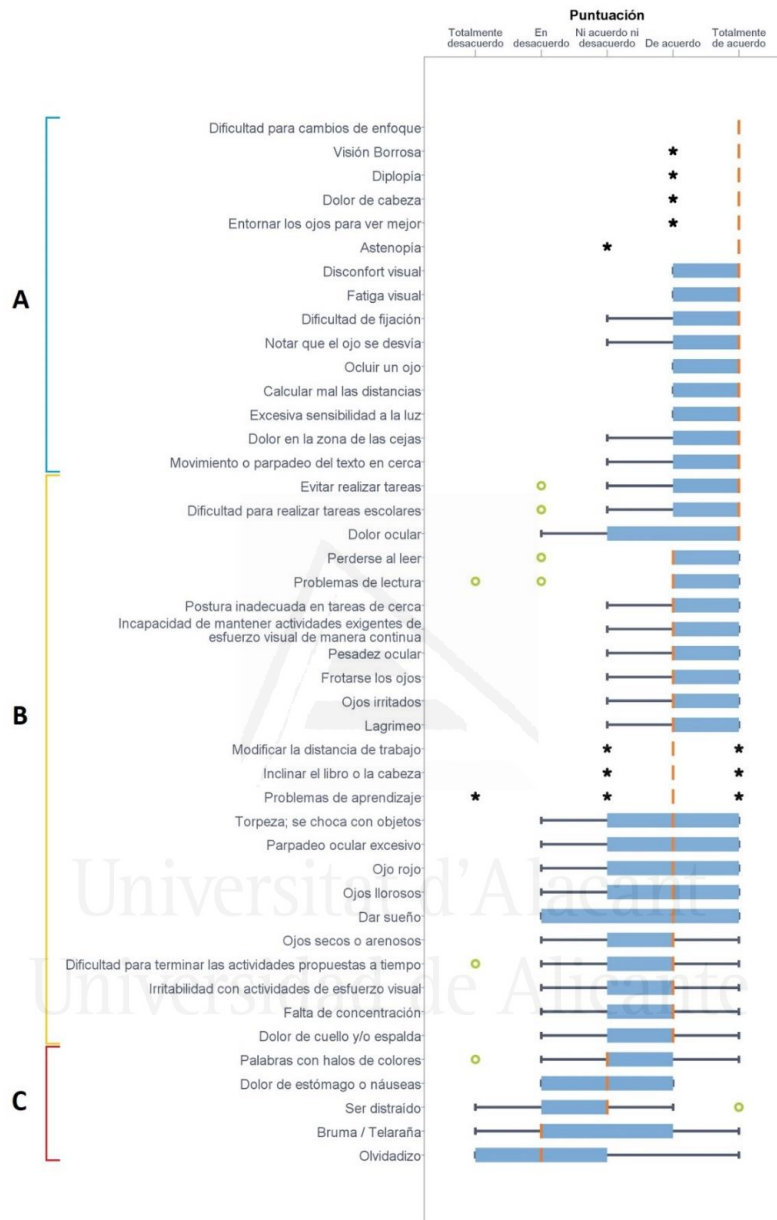


Figura 6. Resultados de la ronda 3. Diagrama de caja y bigotes que muestra el grado de acuerdo de los expertos con respecto a cada uno de los 44 síntomas en un cuestionario de síntomas visuales. A: Síntomas para los cuales los expertos estuvieron de acuerdo en su inclusión en un cuestionario. B: Síntomas con discrepancias con su aparición en un cuestionario. C: Síntomas para los cuales los expertos estuvieron de acuerdo en que no aparecieran en un cuestionario.

De los 15 síntomas que los expertos consideraron que debían aparecer en un cuestionario (zona A de la Figura 6), 11 de ellos figuran en la mayoría de los cuestionarios publicados en la literatura científica.^{34,128-138} Los 4 síntomas que no aparecen son *astenopía*, *dificultad de fijación*, *dolor en la zona de las cejas* y *notar que un ojo se desvía*. Este hallazgo muestra la utilidad de la metodología Delphi, ya que se pueden caracterizar y localizar síntomas comúnmente reportados en la práctica clínica pero que no están considerados por la literatura existente.

De los cinco síntomas que los expertos consideraron que no debían estar en un cuestionario (zona C de la Figura 6), dos de ellos fueron incluidos por el panel de expertos, *bruma / telaraña* y *olvidadizo*. El síntoma *bruma / telaraña* no aparece en ningún cuestionario, y el otro síntoma, *olvidadizo*, se encuentra en el cuestionario 19-item COVID-QOL.¹³⁷ Los otros tres síntomas (*dolor de estómago o náusea*, *palabras con halos de colores* y *ser distraído*) aparecen en varios cuestionarios. El síntoma de *dolor de estómago o náusea* está incluido en el cuestionario CISS inicial de 16 ítems,¹³⁵ aunque posteriormente en las versiones de validación del cuestionario CISS para niños¹³¹ y adultos¹³² ese síntoma se excluyó. El síntoma de *palabras con halos de colores* está incluido en el cuestionario de Adler¹³³ y *ser distraído* aparece en el 19-item COVID-QOL.¹³⁷ Sin embargo, el panel de expertos consideró que, si bien estos cuatro síntomas (*dolor de estómago o náusea*, *olvidadizo*, *palabras con halos de colores* y *ser distraído*) se han utilizado en la literatura científica, tienen poca importancia en el entorno clínico y, por lo tanto, deberían omitirse en un cuestionario. Este hallazgo muestra que algunos de los síntomas considerados por la literatura existente no son importantes para los expertos como para incluirlos en un cuestionario de síntomas.

Para los 24 síntomas restantes (zona B de la Figura 6), los expertos tenían discrepancias con respecto a su inclusión en un cuestionario, por lo que deben analizarse cuidadosamente para su inclusión en un futuro cuestionario. Podrían ser

considerados inicialmente a la hora de desarrollar un cuestionario y los métodos de validación psicométrica adecuados ayudarían a decidir si incluirlos o no. Sobre estos 24 síntomas hubo diferentes opiniones en función de que los expertos ejercieran en una clínica optométrica o en un establecimiento sanitario de óptica. Los expertos que trabajaban en establecimientos sanitarios de óptica dieron importancia a los siguientes 10 síntomas: *dar sueño, dificultad para terminar las tareas propuestas a tiempo, dolor de cuello y/o espalda, dolor ocular, ojos irritados, ojos llorosos, ojos rojos, ojos secos o arenosos, parpadeo ocular excesivo y torpeza / se choca con objetos*. Sin embargo, los expertos que desarrollaban su trabajo en clínicas optométricas no los consideraban tan importantes. Estas discrepancias podrían explicarse por el diferente tipo de pacientes atendidos en cada centro, aunque este hecho no tiene por qué afectar para considerar o no dichos síntomas.

Al analizar las tres rondas en común, se puede observar que existe poca variación a largo del desarrollo de la metodología Delphi en cuanto a la importancia que los expertos le dan a los síntomas. Los 15 síntomas que los expertos coincidieron en que debían aparecer en un cuestionario también fueron considerados importantes en la primera y en la segunda ronda. Y algo similar ocurre con los cinco síntomas que los expertos consideraron que no deberían aparecer en un cuestionario. Es decir, los síntomas *dolor de estómago o náusea, palabras con halos de colores y ser distraído* tuvieron puntuaciones bajas en la primera y segunda ronda para todas las disfunciones (aunque el síntoma *dolor de estómago o náusea* obtuvo una mejor puntuación para las anomalías binoculares). Los otros dos síntomas, *bruma / telaraña* y *olvidadizo* no estaban en la primera ronda y fueron agregados por los expertos en la segunda ronda. En el caso del *olvidadizo*, los expertos coincidieron en que ese síntoma no debe aparecer en un cuestionario y, de hecho, no fue importante ni en la primera ni en la segunda ronda. Sin embargo, se da una situación particular para el síntoma *bruma /*

telaraña, para el cual los expertos consideraron que no era importante estar en un cuestionario y dieron puntuaciones bajas para las disfunciones acomodativas y binoculares, pero una puntuación aceptable para las anomalías refractivas. La pequeña variación entre rondas muestra que, desde el principio, los expertos tenían claro qué síntomas podían ser los más importantes. Como se ha podido observar, hubo variaciones mínimas, situación habitual cuando se realiza una metodología Delphi. De hecho, la importancia de este método radica en que los expertos pueden cambiar impresiones de forma anónima y reflexionar sobre su opinión, pudiendo modificarla a lo largo del proceso. En cualquier caso, esta investigación ha demostrado que los expertos tenían clara su posición sobre varios síntomas, con discrepancias para otros síntomas.

Una de las limitaciones de esta primera parte de la investigación es que los expertos pudieran haberse sentido influenciados por la opinión de otros, lo que sesgaría los resultados. Para evitar esta situación, el grupo coordinador no mostró durante el desarrollo del Delphi las identidades de los expertos. Otra de las limitaciones podría estar relacionada con el hecho de no haber podido contar con optometristas de toda la zona geográfica de España, al haber declinado algunos de ellos la invitación. Aunque en la elección inicial de expertos se intentaba cubrir todas las áreas geográficas nacionales posibles, hubo tres expertos que no quisieron participar. En cualquier caso, el área geográfica que finalmente se abarcó fue lo suficientemente extensa como para que los resultados del estudio no se vieran afectados. De hecho, una de las fortalezas de este estudio es que se obtuvo una tasa de respuesta del 100% por parte de los expertos participantes, ya que todos respondieron a las tres rondas de la metodología Delphi.

En resumen, la metodología Delphi resultó útil para identificar los síntomas relacionados con las disfunciones visuales, tanto en pacientes de clínicas optométricas como de establecimientos sanitarios de óptica, por lo que estos resultados pueden

considerarse como el punto de partida para desarrollar un cuestionario que facilite la recogida y clasificación de estos síntomas.



Universitat d'Alacant
Universidad de Alicante

1.5.2. Análisis Rasch para el desarrollo y reducción del Symptom Questionnaire for Visual Dysfunctions (SQVD)

En la metodología Delphi realizada, los expertos señalaron 39 categorías de síntomas que podían aparecer en un cuestionario (15 en las que existía acuerdo y 24 en las que había cierta controversia). Con estas categorías se diseñó un instrumento de 47 ítems en el que se mostraban diferentes formas de preguntar por la sintomatología del paciente. Con este cuestionario inicial se realizó un pretest cognitivo con el fin de verificar que la población a estudiar comprendía las preguntas y las opciones de respuesta propuestas, además de ser capaz de responder de manera significativa.¹⁵⁷

Este instrumento preliminar de 47 preguntas fue evaluado en una muestra clínica de 30 pacientes de una clínica optométrica, con edades entre 18 y 71 años (edad media 29.83 ± 11.07 años). Además, se realizó una entrevista semiestructurada a estos pacientes donde se les preguntó sobre la longitud del cuestionario, si se entendía bien la forma de responder a las preguntas, si se entendía bien el modo de representar la frecuencia de la respuesta, o si había alguna pregunta que no habían comprendido (especificando cuál). Los pacientes pudieron opinar también sobre la forma de preguntar por determinados síntomas o cuál era la mejor opción de valorar la frecuencia.

Considerando las respuestas de los pacientes que participaron en el pretest cognitivo, se realizó una primera reducción cualitativa de ítems, eliminando preguntas repetidas y realizando los cambios sugeridos por los pacientes. Se obtuvo de esta manera un cuestionario piloto de 33 ítems, denominado Symptom Questionnaire for Visual Dysfunctions (SQVD), en el cual se distribuyeron de forma aleatoria los ítems resultantes del pretest cognitivo. Este fue el punto de partida del segundo artículo de esta tesis.

La ventaja de trabajar con un cuestionario piloto es que permite eliminar elementos no válidos del instrumento, poco fiables o que discriminan de manera deficiente, como paso previo a la obtención del cuestionario definitivo. De este modo, en este segundo artículo se utilizó el análisis Rasch para realizar una reducción de los ítems del cuestionario piloto de 33 preguntas.

En el cuestionario piloto de 33 preguntas se redactaron los ítems como preguntas de opción múltiple con una escala Likert de 4 categorías de respuesta que medía la frecuencia del ítem. El significado de cada opción de respuesta, el cual se reflejó en las instrucciones del cuestionario, fue el siguiente:

- No: el síntoma nunca ocurre.
- Ocasionalmente: el síntoma se presenta al menos una vez cada 15 días.
- A menudo: el síntoma se presenta una o dos veces por semana.
- Casi siempre: el síntoma ocurre casi todos los días.

Este cuestionario piloto de 33 preguntas se distribuyó en una muestra poblacional consecutiva de 125 pacientes de una clínica optométrica, con edades comprendidas entre los 15 y los 84 años (41.72 ± 14.92 años). Se incluyeron sujetos sin ningún tipo de patología ocular. De la muestra clínica, 74 personas (59.2%) eran mujeres y 51 (40.8%) hombres, siendo 57 pacientes (45.6%) présbitas. De todos ellos, 95 pacientes presentaban disfunciones refractivas, 18 disfunciones binoculares, 7 anomalías acomodativas y 5 sujetos no presentaban ninguna anomalía visual. Todos los pacientes firmaron el consentimiento informado.

Con el fin de valorar si existían ítems candidatos a ser eliminados y analizar psicométricamente las características del cuestionario resultante, se realizó un análisis Rasch mediante el Andrich Rating Scale Model (ARSM),¹⁴¹ ya que el *likelihood ratio statistic* fue no significativo ($p > 0.05$). Este análisis Rasch incluyó el estudio de las

curvas de probabilidad de respuesta, la reducción de ítems (mediante el ajuste de los ítems al modelo usando los estadísticos infit y outfit y la dependencia local a través del estadístico Yen-Q3), el análisis del funcionamiento diferencial de los ítems (differential item functioning, DIF), la fiabilidad de la separación de las medidas de las personas y de los ítems, la unidimensionalidad mediante el análisis de componentes principales de los residuales y la medida del alineamiento entre ítems y personas (targeting). Utilizando el análisis Rasch también se calculó la tabla de transformación de puntuación ordinal a intervalos. El software utilizado fue Winsteps versión 4.8.1.

1.5.2.1. Análisis de las curvas de probabilidad de respuesta

El primer paso realizado fue el análisis de las curvas de probabilidad de respuesta con el fin de verificar si la calibración de las categorías del cuestionario aumentaba de manera ordenada (Figura 7). En primer lugar, en la Figura 7a, se puede observar cómo el umbral de las categorías de respuesta 1 y 2 y el umbral de las categorías 2 y 3 se corresponden con el mismo valor de la medida de la persona en relación con la dificultad del ítem. Esto implica que las categorías de respuesta no aumentan adecuadamente y los umbrales están desordenados. Al no estar ordenadas las categorías de respuesta, se examinaron los umbrales de Andrich (Tabla 6). Los resultados muestran que, para las cuatro categorías iniciales (Tabla 6A), los umbrales están separados menos de 1.4 logits, lo que unido a la interpretación de la Figura 7a, implica que es recomendable modificar las categorías de respuesta.^{139,161} Cuando se colapsan las categorías 1 y 2 (Tabla 6B) se observa que existe una diferencia entre categorías de 2.68 logits, y cuando se colapsan las categorías 2 y 3 esa diferencia es de 0.74 logits (Tabla 6C). A raíz de este resultado, las categorías iniciales 1 (Ocasionalmente) y 2 (A menudo) se colapsaron en la categoría Ocasionalmente / A

menudo, dando como resultado tres opciones de respuesta de categoría: 0 (No), 1 (Ocasionalmente / A menudo) y 2 (Casi siempre).

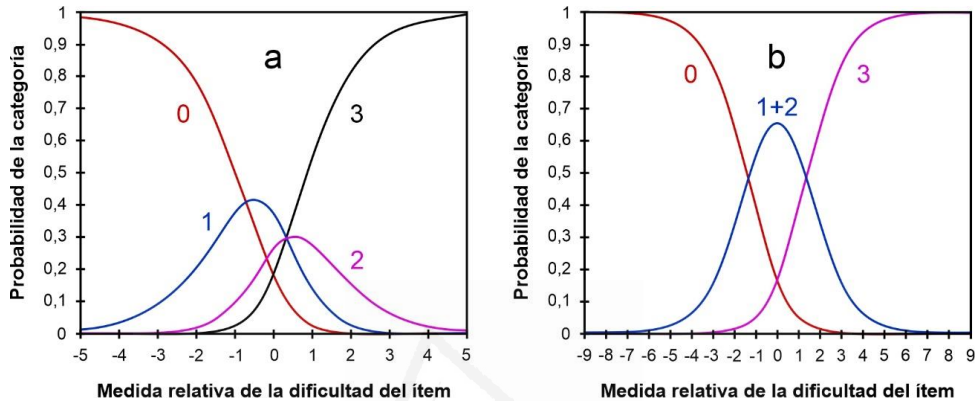


Figura 7. La Figura 7a muestra el análisis de las curvas de probabilidad de respuesta para la versión de 33 ítems del SQVD con cuatro categorías de respuesta. Cada curva en el gráfico representa una categoría de respuesta (No = 0; Ocasionalmente = 1; A menudo = 2; Casi siempre = 3). El punto donde se superponen dos curvas adyacentes es el umbral. En esta intersección, existe la misma probabilidad de elegir una categoría u otra. La Figura 7b muestra el análisis de las curvas de probabilidad de respuesta para la versión de 33 ítems del SQVD colapsando las categorías 1 (Ocasionalmente) y 2 (A menudo).

El resultado de este colapso de categorías se puede ver en la Figura 7b, y se aprecia que en la versión piloto de 33 ítems del SQVD con tres categorías de respuesta los umbrales están ordenados y con separaciones superiores a 1.4 logits entre ellos. Esta situación de colapso de categorías también ha sido utilizada por varios cuestionarios relacionados con la visión.^{119,120,122}

	Categoría	Umbral de Andrich
A	0	Ninguno
	1	-0.67
	2	0.25
	3	0.42
B	0	Ninguno
	1+2	-1.34
	3	1.34
C	0	Ninguno
	1	-0.37
	2+3	0.37

Tabla 6. Valores de los umbrales de Andrich (en logits) para la versión piloto del SQVD con 33 ítems, con sus cuatro categorías originales y categorías colapsadas. A: resultados para la versión piloto de 33 ítems con las cuatro categorías iniciales de respuesta (0-1-2-3). B: resultados al colapsar las categorías 1 y 2. C: resultados al colapsar las categorías 2 y 3.

1.5.2.2. Reducción psicométrica de los ítems

Una vez colapsadas las categorías de respuesta del instrumento, se comprobó si debía realizarse una reducción psicométrica de los ítems mediante el ajuste de los ítems al modelo usando los estadísticos infit y outfit, y mediante el análisis de la dependencia local.

Siguiendo los criterios para la reducción de ítems basados en la literatura científica^{123,139,159,162} y descritos anteriormente en la sección **1.1.5. Desarrollo y validación de cuestionarios**, la versión piloto del SQVD se redujo de 33 preguntas a 16 preguntas. En esta reducción se eliminaron 17 ítems para dar lugar a una versión del SQVD de 16 ítems.

Con esta versión de 16 ítems del cuestionario, se realizó el análisis de la dependencia local de los ítems a través del estadístico Yen-Q3 para comprobar si había

que eliminar más preguntas. Como la correlación promedio de la matriz de correlación residual fue de 0.12, los valores de correlación mayores a 0.32 (valor que corresponde a $0.12 + 0.2$) indicaban que existía dependencia local, lo cual implicaba que había que eliminar estos ítems del cuestionario. La Tabla 7 muestra las parejas de ítems para las que se obtuvieron valores de correlación residual mayores a este valor de 0.32, donde además se han especificado los ítems que había que eliminar o retener en función de los criterios de dependencia local para la reducción de ítems. Como se comentó en la sección **1.1.5. Desarrollo y validación de cuestionarios**, se conservaron los ítems con un valor de outfit MNSQ más cercano a 1. Una vez aplicado este criterio, hubo que eliminar los ítems 28 y 31, por lo que el cuestionario final quedó compuesto por 14 ítems. Se volvió a analizar si existía dependencia local en estos 14 ítems, y se comprobó que no había dependencia.

Correlación residual	Ítem	Outfit MNSQ	Resultado	Ítem	Outfit MNSQ	Resultado
0.41	31	0.88	Eliminado	33	1.01	Retenido
0.37	28	0.92	Eliminado	33	1.01	Retenido
0.34	25	0.95	Retenido	31	0.88	Eliminado
0.33	28	0.92	Eliminado	31	0.88	Eliminado

Tabla 7. Análisis de dependencia local mediante el estadístico Yen-Q3. (MNSQ: cuadrático medio).

Por tanto, en total, se eliminaron 19 ítems de la versión piloto de 33 ítems del SQVD. Estos 19 ítems eliminados estaban relacionados con problemas de lectura, fatiga visual, diplopía, problemas oculares, visión borrosa y problemas posturales. Así, la versión final del SQVD se compuso de 14 ítems, una cantidad de ítems similares a los cuestionarios mostrados en la literatura científica relacionados con síntomas

visuales^{34,130-132} Estos 14 ítems estaban relacionados con visión borrosa, problemas de visión binocular, irritación ocular, dificultad de concentración, problemas de lectura y dolor de cabeza.

1.5.2.3. Ajuste de los ítems al modelo usando los estadísticos infit y outfit

Con la versión de 14 ítems del SQVD, se volvió a realizar el análisis del ajuste de los ítems al modelo usando los estadísticos infit y outfit MNSQ. En la Tabla 8 se pueden ver estos valores para cada uno de los ítems. Como se observa, para el infit MNSQ se obtuvieron valores entre 0.78 y 1.19 y para el outfit MNSQ entre 0.74 y 1.24. De estos resultados se desprende que los 14 ítems del cuestionario demostraron individualmente unos valores de estadísticos que se ajustaban adecuadamente al modelo Rasch. En cuanto a los valores medios de los estadísticos globales de infit y outfit MNSQ del cuestionario final, tanto para las personas como los ítems, dichos resultados se muestran en la Tabla 9. Como puede observarse, tanto los resultados de infit y outfit MNSQ para las personas (infit 0.99 y outfit 0.99) como para los ítems (infit 0.99 y outfit 0.99) también se encuentran del rango recomendado por la literatura científica.^{123,139,159,161,190} Estos resultados implican por tanto la existencia de un adecuado ajuste al modelo Rasch.

Ítem	Infit MNSQ	Outfit MNSQ
1	0.99	0.96
2	1.04	0.97
6	0.98	0.91
8	0.98	0.95
12	1.01	1.00
14	0.96	1.06
15	1.03	1.06
17	0.78	0.74
19	1.19	1.24
23	0.86	0.85
25	1.10	0.99
26	0.98	1.07
32	0.84	0.95
33	1.18	1.13

Tabla 8. Resultados de los estadísticos de ajuste Infit y Outfit MNSQ para los 14 ítems del cuestionario SQVD (125 pacientes). (MNSQ: cuadrático medio).

Personas		Ítems	
Infit MNSQ	Outfit MNSQ	Infit MNSQ	Outfit MNSQ
0.99	0.99	0.99	0.99

Tabla 9. Resumen de los estadísticos de ajuste global para las habilidades de las personas y dificultad de los ítems para la versión de 14 ítems del SQVD (125 pacientes). (MNSQ: cuadrático medio).

1.5.2.4. Análisis del funcionamiento diferencial de los ítems (DIF)

El análisis del funcionamiento diferencial de los ítems se realizó dividiendo la muestra por sexo y por presbicia (présbitas y no présbitas, considerando la presbicia cuando el paciente necesitaba adición). Los resultados del análisis DIF (Tabla 10) muestran que no hubo diferencias estadísticamente significativas ($p > 0.05$) con respecto al sexo para todos los ítems (< 1 logits). En cambio, en el análisis de la presbicia, no hubo diferencias estadísticamente significativas para los ítems (< 1 logits) excepto para el ítem 33. Este ítem está relacionado con la visión borrosa de cerca, por lo que se podría argumentar que este síntoma es más específico de una población présbita, y es por ello que este ítem no se consideró que debía ser eliminado del cuestionario. En cualquier caso, los resultados del DIF permiten asumir que el modelo y el conjunto de parámetros de los ítems son similares para todos los grupos comparables.

Ítem	Contraste DIF para sexo	Contraste DIF para presbicia
1	0.37	0.33
2	0.32	0.13
6	0.29	0.49
8	0.00	0.08
12	0.53	0.10
14	0.33	0.89
15	0.37	0.71
17	0.37	0.00
19	0.19	0.53
23	0.19	0.05
25	0.36	0.21
26	0.24	0.13
32	0.13	0.00
33	0.07	1.71*

Tabla 10. Resultados del análisis DIF para los 14 ítems del SQVD (125 pacientes) según las variables de sexo y presbicia (DIF: Funcionamiento diferencial de los ítems). * $p < 0.05$.

1.5.2.5. Fiabilidad de la separación de las medidas de las personas y de los ítems

Los resultados de los índices de fiabilidad (Reliability, Person separation) de las personas y de los ítems se muestran en la Tabla 11. Como se observa, la fiabilidad (Reliability) de las personas fue de 0.81 (Person separation: 2.11), lo que implica una buena consistencia interna para el SQVD sobre la muestra seleccionada. Dicho resultado indica que el cuestionario puede estratificar la población en al menos 3 grupos según el rasgo latente medido. Sin embargo, la fiabilidad (Reliability) obtenida para los ítems fue de 0.80 (Item separation: 2.06), un valor bajo que muestra que sería necesaria una muestra mayor para demostrar la jerarquía de dificultad de los ítems de la escala.

Fiabilidad (Separation index)	
Fiabilidad de las personas (Separation index)	Fiabilidad de los ítems (Separation index)
0.81 (2.11)	0.80 (2.06)

Tabla 11. Resultados de los índices de fiabilidad (Reliability, Separation Index) para las personas y los ítems del SQVD (125 pacientes).

1.5.2.6. Unidimensionalidad

La unidimensionalidad de la escala del SQVD se demostró mediante el análisis de componentes principales (PCA), donde, para la magnitud del primer contraste del residual, se obtuvo un valor de 1.908 eigenvalue, resultado inferior a 2 tal y como se recomienda en la literatura científica para mostrar la unidimensionalidad.¹⁴¹ Dicho resultado implica por tanto que se puede asegurar que la versión de 14 ítems del SQVD

es un instrumento unidimensional, lo que significa que el SQVD evalúa un único constructo, es decir, todos los ítems miden lo mismo.

1.5.2.7. Medida del alineamiento entre ítems y personas (targeting)

El análisis de la medida del alineamiento entre los ítems y las personas, conocido como targeting, se puede apreciar en la Figura 8. En dicha figura, correspondiente al mapa de personas-ítem, la columna de la izquierda indica la habilidad de las personas (en logits), donde los sujetos con mayor habilidad se muestran en la parte superior de la figura. La habilidad media de las personas (mostrada en la figura como la M de la izquierda) tiene un valor de -1.59 logits. La columna de la derecha indica la dificultad del ítem, para la cual la media (M derecha) es siempre 0. En consecuencia, los ítems más difíciles se encuentran en la parte superior de la figura. En este caso, el targeting del SQVD cuyo valor es de -1.59 logits, indica un targeting pobre, revelando lo que se conoce como efecto suelo. La explicación se puede ver claramente en la Figura 8 donde existe un número importante de pacientes que se encuentran en la parte inferior de la figura. Esto podría explicarse porque, en la muestra actual, la sintomatología referida por los sujetos es leve, o lo que es lo mismo, hay un mayor número de sujetos con menor sintomatología. En la literatura científica, algunos autores han mostrado que esta situación es común en los cuestionarios relacionados con la sintomatología, ya que puede deberse a que muchos pacientes pueden no presentar síntomas^{120,121} o a que existe una tendencia de los sujetos a no decir que los presentan.¹¹⁹ Estas condiciones son aplicables al SQVD, por lo que el resultado del targeting obtenido en esta investigación puede considerarse aceptable.

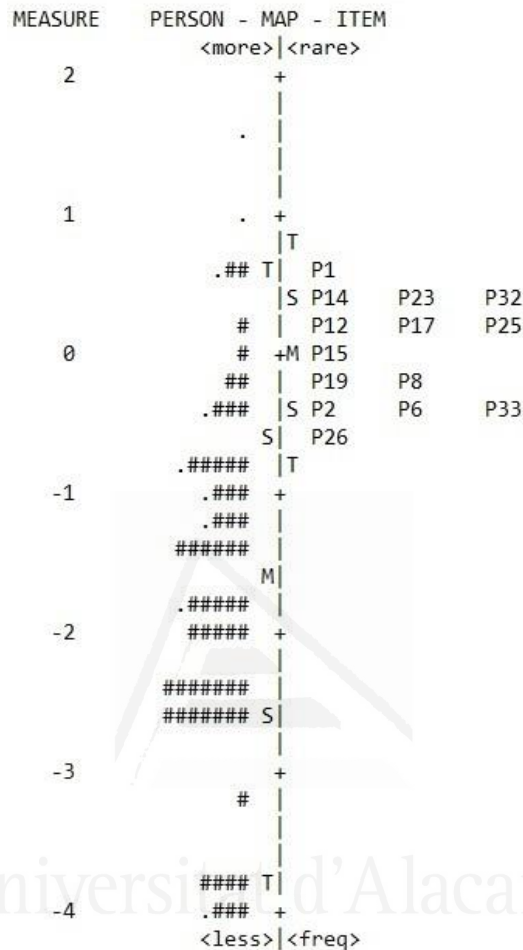


Figura 8. Mapa personas-ítem para el SQVD. Los sujetos están representados a la izquierda de la línea discontinua con el símbolo "#" (representa a 2 sujetos) y "." (representa a 1 sujeto). A la derecha de la línea discontinua se muestran los ítems de la versión SQVD de 14 ítems con su número (P y Número del ítem). M indica el valor medio (a la izquierda la habilidad de las personas y a la derecha la dificultad de los ítems). S muestra una desviación estándar de la media y T muestra dos desviaciones estándar. En la parte superior de la figura se representa una mayor habilidad para las personas (mayor frecuencia de síntomas) y una mayor dificultad para los ítems.

1.5.2.8. Tabla de transformación de las puntuaciones

Los resultados de la tabla de transformación se muestran en la Tabla 12, donde se pueden ver la escala ordinal y las puntuaciones correspondientes a nivel de intervalo.¹⁴¹ Las puntuaciones ordinales del cuestionario oscilan entre 0 y 28 y se incluyen las puntuaciones de intervalo correspondientes en logits. Las puntuaciones de intervalo en unidades logit se transformaron en el rango de la escala ordinal (0-28).

Con esta tabla de conversión, se ha establecido que cuando se utiliza un cuestionario, se puede aumentar la precisión de los resultados puesto que la conversión a puntuaciones de intervalo facilita el uso de estadística paramétrica. Por ese motivo, esta tabla puede ser útil para fines clínicos, por ejemplo, cuando se quiere informar de cambios de la variable de medida, ya que con una escala de intervalos iguales se pueden detectar variaciones en la puntuación del cuestionario. Sin embargo, hay que tener en cuenta que estos valores de la tabla de transformación solamente se pueden utilizar cuando las personas al contestar al cuestionario, contestan a todas las preguntas y por tanto no existen preguntas sin contestar.

Medida en escala Ordinal	Medida de intervalo	
	Puntuación cruda	Logit
0	-5.19	0.00
1	-3.93	3.39
2	-3.17	5.45
3	-2.69	6.75
4	-2.32	7.74
5	-2.01	8.57
6	-1.74	9.30
7	-1.50	9.96
8	-1.27	10.59
9	-1.04	11.18
10	-0.83	11.76
11	-0.62	12.33
12	-0.41	12.89
13	-0.21	13.45
14	0.00	14.00
15	0.21	14.55
16	0.41	15.11
17	0.62	15.67
18	0.83	16.24
19	1.04	16.82
20	1.27	17.41
21	1.50	18.04
22	1.74	18.71
23	2.01	19.43
24	2.32	20.26
25	2.69	21.25
26	3.17	22.55
27	3.93	24.61
28	5.19	28.00

Tabla 12. Conversión de la puntuación del SQVD (125 pacientes) sin procesar (0 - 28) a una escala de intervalo en unidades logit y utilizando las métricas de la escala original.

A la hora de comparar este cuestionario SQVD con otros cuestionarios existentes que estén relacionados con síntomas visuales, este análisis es difícil de realizar ya que sus objetivos son diferentes. Así, el cuestionario de Conlon¹³⁰ es un instrumento para medir los síntomas del disconfort visual, pero esto es diferente a la sintomatología asociada a las disfunciones visuales. El disconfort visual puede deberse a diferentes causas, como por ejemplo el deslumbramiento producido por una fuente de luz, por lo que no es comparable al SQVD. Asimismo, el CISS^{34,131,132} es un cuestionario específico sólo para una anomalía binocular, la insuficiencia de convergencia, por lo que no puede aplicarse a las demás anomalías visuales, ya que un cuestionario solo puede ser utilizado para la condición para la cual fue desarrollado y validado.^{181,182} Por tanto, no se pueden comparar las categorías de síntomas del CISS y del SQVD. Además, el cuestionario CISS no ha sido validado con el modelo Rasch por lo que no se puede conocer si realmente las preguntas del cuestionario son o no adecuadas para lo que se pretende medir.

Esta segunda parte de la investigación que conforma esta tesis doctoral también puede haber tenido algunas limitaciones. A pesar de que la versión piloto de 33 ítems del SQVD se distribuyó a 125 sujetos de una consulta optométrica, el tamaño de la muestra puede considerarse pequeño. Varios autores¹⁹¹ han sugerido que el número de sujetos necesarios para analizar un cuestionario podría variar entre 4 y 10 sujetos por ítem, con un mínimo de 100 sujetos. Sin embargo, otros autores¹⁸² han demostrado que no existen reglas absolutas para el cálculo del tamaño muestral necesario para validar un cuestionario.

En cualquier caso, también se pueden remarcar varios puntos fuertes. Uno de ellos está relacionado con las características poblacionales de la muestra. Desde un punto de vista clínico, un cuestionario como el SQVD tiene beneficios prácticos para fines clínicos ya que puede ser una ayuda para calibrar los síntomas de los pacientes

que acuden a una clínica optométrica. Debido a que la población clínica debe ser la población diana de este instrumento, es una fortaleza que su desarrollo se haya realizado utilizando una muestra clínica. Además, el punto de partida del estudio es también otra fortaleza a considerar. El meticuloso proceso desarrollado para determinar qué síntomas debían incluirse en un cuestionario de síntomas visuales (el método Delphi anteriormente utilizado),¹⁹² así como el procedimiento desarrollado para diseñar la escala inicial y el pretest cognitivo previo, son otro de los puntos fuertes que contribuyen a la solidez del estudio.

En resumen, esta segunda parte de la tesis doctoral muestra que la versión del SQVD de 14 ítems es un instrumento cuyos ítems se ajustan adecuadamente al modelo Rasch, sin dependencia local o DIF estadísticamente significativo para las categorías de sexo y presbicia, y con una adecuada fiabilidad y unidimensionalidad.

1.5.3. Validación del Symptom Questionnaire for Visual Dysfunctions (SQVD): un cuestionario para evaluar los síntomas de cualquier tipo de disfunción visual

Una vez realizada la reducción de ítems y obtenido el cuestionario SQVD de 14 ítems, así como realizado un primer análisis psicométrico del SQVD, es necesario analizar y confirmar los resultados de la versión final del cuestionario en una muestra clínica de mayor tamaño. Para ello, se realizó el proceso de validación del SQVD en una muestra clínica de pacientes con cualquier tipo de anomalía visual, analizando no solo sus propiedades psicométricas, sino la exactitud diagnóstica, la validez y la repetibilidad del instrumento.

Para poder llevar a cabo este proceso de validación, se distribuyó el cuestionario SQVD a una muestra clínica de una consulta optométrica. De este modo, todos aquellos sujetos con edad a partir de 14 años que acudían consecutivamente a realizarse un examen visual en una clínica optométrica eran candidatos a incluirse en esta tercera parte de la investigación. A cada paciente se le realizaba un examen optométrico completo con el fin de determinar la existencia de algún tipo de anomalía visual.⁴⁹ En dicho examen optométrico, se recogían además los síntomas que los pacientes referían durante la historia clínica. Una vez obtenida toda esta información, se realizaba el diagnóstico de cualquier tipo de disfunción visual. Para dicho diagnóstico se siguieron los criterios utilizados por García-Muñoz et al.⁴⁹ De este modo, una vez diagnosticado cualquier tipo de disfunción visual, todos aquellos sujetos que presentaban cualquier tipo de patología ocular, cirugía refractiva, ojo seco o que tomaban algún tipo de medicación que pudiera alterar la función visual fueron excluidos del estudio.

De este modo, siguiendo dichos criterios de inclusión, el estudio incluyó a una muestra de 306 pacientes con edades comprendidas entre 14 y 87 años (edad media: 38.38 ± 17.29 años), de los que 184 (60.1%) eran mujeres y 120 (39.2%) eran présbitas.

De la muestra analizada de 306 sujetos, 204 pacientes (66.7%) presentaban algún tipo de disfunción visual (132 disfunciones refractivas y 72 sujetos con anomalías acomodativas y / o binoculares). En particular, 14 pacientes tenían disfunciones acomodativas (1 insuficiencia de acomodación y 13 excesos de acomodación) y 50 sujetos tenían anomalías binoculares (17 insuficiencias de convergencia, 10 excesos de convergencia, 4 insuficiencias de divergencia, 5 endoforias básicas, 4 exoforias básicas, 1 hiperforia, 3 ambliopías y 6 estrabismos). Además, 8 sujetos tenían disfunciones acomodativas unidas a anomalías binoculares (4 con exceso de acomodación + insuficiencia de convergencia; 1 con exceso de acomodación + exceso de convergencia; 1 con exceso de acomodación + insuficiencia de divergencia; 2 con insuficiencia de acomodación + insuficiencia de convergencia). Los 102 sujetos restantes (33.3%) por tanto no tenían anomalías visuales.

Para poder realizar el reclutamiento de sujetos de este estudio, se obtuvo el consentimiento informado de todos ellos después de explicarles en qué consistía la investigación. Si el paciente era menor de 18 años, eran los padres o tutores legales de los menores los que tenían que aceptar la participación del menor en la investigación y firmar el consentimiento informado, además de que el menor también debía dar su consentimiento para participar. Todos los pacientes del estudio, los 306 sujetos, completaron el cuestionario SQVD. El cuestionario se auto administraba antes del examen optométrico mediante una copia impresa en la que los pacientes podían leer cada pregunta con sus correspondientes opciones de respuesta. Uno de los investigadores estaba en todo momento con cada paciente para apoyarlos en caso de que tuvieran alguna duda durante la cumplimentación del cuestionario. Este investigador era una persona diferente al examinador que realizó el examen visual para evitar sesgos en el estudio.

El SQVD se muestra en el Anexo 3. El cuestionario consta de 14 ítems que se relacionan con visión borrosa, problemas binoculares, irritación ocular, dolor de cabeza, dificultades de concentración, problemas de lectura y problemas posturales. Cada ítem consta de tres opciones de categoría de respuesta mediante una escala Likert, midiendo la frecuencia de los síntomas. El significado de cada respuesta, que se describe en las instrucciones del instrumento, es el siguiente:

- No: el síntoma nunca ocurre.
- Ocasionalmente / A menudo: el síntoma ocurre al menos una vez cada quince días / una o dos veces por semana.
- Casi siempre: el síntoma ocurre casi todos los días.

El cuestionario SQVD se puntúa dando una puntuación numérica para cada respuesta, de modo que la puntuación total se registra de la siguiente manera: no: 0, ocasionalmente / a menudo: 1, y casi siempre: 2. La puntuación total se obtiene luego sumando los puntos de los 14 ítems, que van de 0 a 28 puntos.

Una vez obtenidos todos los cuestionarios de las 306 personas de esta investigación, se realizaron todos los pasos que se han explicados anteriormente en la sección **1.1.5. Desarrollo y validación de cuestionarios** y que la evidencia científica ha mostrado que son los necesarios y adecuados para realizar la validación del instrumento SQVD.

El análisis Rasch se realizó utilizando el método ARSM¹⁶⁰ mediante el software Winsteps versión 4.8.1. Los parámetros analizados fueron las curvas de probabilidad de respuesta, el ajuste de los ítems al modelo usando los estadísticos infit y outfit, el análisis del funcionamiento diferencial de los ítems (DIF), la fiabilidad de la separación de las medidas de las personas y de los ítems, la medida del alineamiento entre ítems y personas (targeting), la dependencia local a través del estadístico Yen-Q3 y la

unidimensionalidad mediante el análisis de componentes principales de los residuales (PCA). Además, se analizó la exactitud diagnóstica mediante las curvas ROC utilizando como Gold-estándar los síntomas de los pacientes descritos en la historia clínica. También se evaluó la validez del SQVD a través de la validez concurrente y la validez de grupos conocidos. Y la repetibilidad del cuestionario se examinó mediante el coeficiente de correlación intraclass (ICC), distribuyéndolo de nuevo a 50 personas una semana después de la primera visita. Estos análisis de validez y repetibilidad se realizaron mediante el software estadístico SPSS 20.0.

1.5.3.1. Análisis de las curvas de probabilidad de respuesta

En la Figura 9 se pueden ver las curvas de probabilidad de respuesta obtenidas para la población analizada en este estudio. Como se puede observar, en este caso sí existe una buena discriminación de categorías, estando ordenadas las categorías de respuesta. Esto se corrobora con los resultados de los umbrales de Andrich (Tabla 13), donde se puede observar que los valores entre la categoría 1 y 2 estaban separados 2.68 logits. Puesto que este valor es superior a 1.4 logits, esto indica que los umbrales estaban ordenados y las categorías de respuesta de ítems tenían la misma probabilidad de ser contestadas por los sujetos.

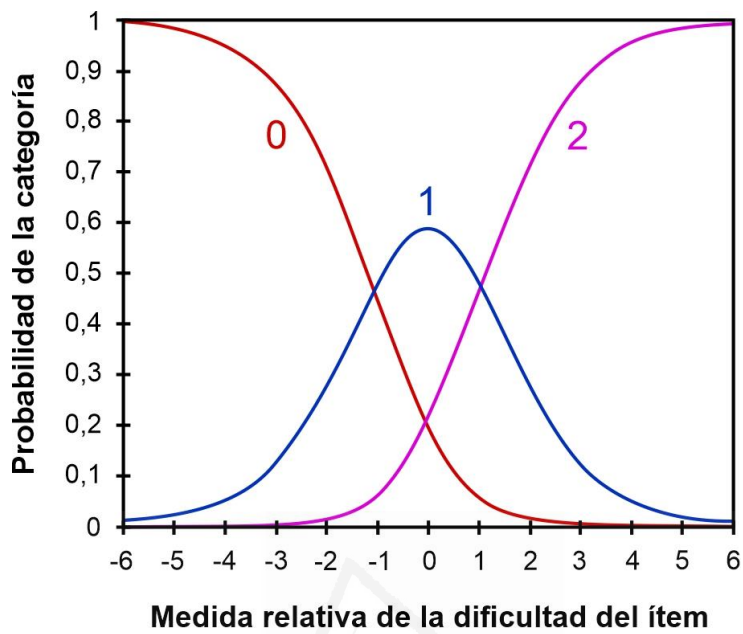


Figura 9. Análisis de las curvas de probabilidad de respuesta (CPC) para el cuestionario SQVD con tres categorías de respuesta (0, 1 y 2).

Categoría	Umbral de Andrich
0	Ninguno
1	-1.34
2	1.34

Tabla 13. Valores de los umbrales de Andrich (en logits) para el SQVD.

1.5.3.2. Ajuste de los ítems al modelo usando los estadísticos infit y outfit

Para analizar el ajuste de los ítems al modelo, se utilizaron los estadísticos infit MNSQ y outfit MNSQ. Los valores de estos estadísticos para cada uno de los 14 ítems del SQVD se muestran en la Tabla 14. Estos valores de los estadísticos de ajuste indicaron que todos los ítems del SQVD se ajustaban al modelo Rasch, ya que todos ellos se encontraban entre los valores de rango sugeridos por la literatura científica (entre 0.70 y 1.30 logits).¹⁴¹ Asimismo, los resultados de los estadísticos de ajuste global infit MNSQ y outfit MNSQ, tanto para las personas como para los ítems (Tabla 15) presentan unos valores medios de infit y outfit MNSQ para las personas (infit 1.01 y outfit 1.00) como para los ítems (infit 1.00 y outfit 1.00). Estos valores indican un muy buen ajuste al modelo Rasch.

Ítem	Infit MNSQ	Outfit MNSQ
1	0.93	0.94
2	1.07	1.06
3	0.93	0.95
4	1.25	1.25
5	0.89	0.98
6	1.00	0.93
7	0.90	0.91
8	0.90	0.87
9	1.14	1.10
10	1.04	1.05
11	1.05	1.14
12	1.15	1.12
13	0.73	0.70
14	0.96	1.03

Tabla 14. Resultados de los estadísticos de ajuste Infit y Outfit MNSQ para los 14 ítems del cuestionario SQVD (306 pacientes). (MNSQ: cuadrático medio).

Personas		Ítems	
Infit MNSQ	Outfit MNSQ	Infit MNSQ	Outfit MNSQ
1.01	1.00	1.00	1.00

Tabla 15. Resumen de los estadísticos de ajuste global para las habilidades de las personas y dificultad de los ítems para el SQVD (306 pacientes). (MNSQ: cuadrático medio).

1.5.3.3. Análisis del funcionamiento diferencial de los ítems (DIF)

Para realizar el análisis DIF, en este caso se tuvieron en cuenta grupos de pacientes en función de tres variables: el sexo (hombre/mujer), las disfunciones (tener o no algún tipo de disfunción visual) y la presbicia (présbitas y no présbitas). En la Tabla 16 se indican los resultados del análisis DIF. Como se puede observar, en general no hubo diferencias estadísticamente significativas para DIF en cuanto al sexo, disfunciones y presbicia (<1 logits). Esto significa que no existe un comportamiento diferente de los ítems para estas variables. Solo dos ítems (ítems 1 y 2 relacionados con dolor de cabeza y visión cercana borrosa respectivamente) mostraron DIF significativo para presbicia. El hecho de que estos dos ítems presentaran DIF para la categoría de presbicia indica que dichos ítems están relacionados con la presbicia, sugiriendo una sintomatología particular de este tipo de población. En consecuencia, se consideró que ambos ítems no debían eliminarse del cuestionario, ya que informan sobre los síntomas de una afección visual importante.

Ítem	Contraste DIF para sexo	Contraste DIF para presbicia	Contraste DIF para disfunción
1	0.25	1.15*	0.20
2	0.02	1.10*	0.26
3	0.00	0.09	0.26
4	0.10	0.49	0.00
5	0.10	0.14	0.35
6	0.15	0.48	0.13
7	0.00	0.24	0.45
8	0.47	0.21	0.36
9	0.58	0.92	0.13
10	0.35	0.08	0.51
11	0.10	0.00	0.10
12	0.52	0.55	0.00
13	0.24	0.60	0.47
14	0.14	0.80	0.31

Tabla 16. Resultados del análisis DIF para los 14 ítems del SQVD (306 pacientes) según las variables de sexo, presbicia y disfunción. (DIF: Funcionamiento diferencial de los ítems). * $p < 0.05$.

1.5.3.4. Fiabilidad de la separación de las medidas de las personas y de los ítems

Los resultados de los índices de fiabilidad (Reliability, Person separation) de las personas y de los ítems se muestran en la Tabla 17. Se observa que la fiabilidad (Reliability) de las personas fue de 0.81 (Person separation: 2.11), lo que implica una buena consistencia interna para el SQVD sobre la muestra seleccionada. Este resultado indica que el cuestionario puede estratificar la población en al menos 3 grupos según el rasgo latente medido. Sin embargo, la fiabilidad (Reliability) obtenida para los ítems fue de 0.85 (Item separation: 2.41), un valor bajo que muestra que sería necesaria una muestra mayor para demostrar la jerarquía de dificultad de los ítems de la escala.

Fiabilidad (Separation index)	
Fiabilidad de las personas (Separation index)	Fiabilidad de los ítems (Separation index)
0.81 (2.11)	0.85 (2.41)

Tabla 17. Resultados de los índices de fiabilidad (Reliability, Separation Index) para las personas y los ítems del SQVD (306 pacientes).

1.5.3.5. Medida del alineamiento entre ítems y personas (targeting)

En cuanto al targeting, la Figura 10 muestra el mapa personas-ítem para el cuestionario SQVD recopilado en la muestra de 306 personas. En esta muestra, el valor medio de la habilidad de las personas fue de -1.49 logits (resultado mostrado en la figura con la letra M de la izquierda). La columna de la derecha indica la dificultad del ítem, para la cual la media es siempre 0 (se muestra con la letra M de la derecha). Este pobre valor del targeting (-1.49 logits) encontrado en este estudio, muestra de nuevo al igual que en el estudio previo del cuestionario, que un número importante de pacientes se situó en la parte inferior de la figura, lo que indica un efecto suelo. Por tanto, este mapa que relaciona a las personas y los ítems está indicando que los ítems están alineados con los pacientes más sintomáticos. Por tanto, los pacientes que tienen menos síntomas visuales no podrán diferenciarse bien por los ítems del cuestionario.

En cualquier caso, hay que tener en cuenta que este resultado también se ha encontrado en otros cuestionarios que evalúan los síntomas de las personas, de modo que este pobre targeting lo que está explicando es que cuando a las personas se les pregunta por sus síntomas, puede que en la muestra analizada ocurra que los pacientes no tengan esos síntomas que se están evaluando^{120,121} o simplemente que a veces los pacientes tienden a no decir que tienen síntomas.¹¹⁹ Por lo tanto, aunque el valor del

targeting para el SQVD no sea del todo bueno, puede considerarse razonable teniendo en cuenta estas consideraciones comentadas. De todos modos, ciertamente sería necesario que en futuros estudios se pudieran analizar muestras de pacientes con niveles más altos de síntomas para mejorar el valor de targeting. De hecho, se puede observar que el targeting obtenido en este estudio es sensiblemente mejor que el hallado en la segunda investigación de esta tesis, en la que se utilizó una muestra menor y probablemente con menor sintomatología.

1.5.3.6. Unidimensionalidad y dependencia local

En el análisis de la unidimensionalidad, el valor de componentes principales de los residuales (PCA) mostró la unidimensionalidad del SQVD, ya que la magnitud del primer contraste del residual obtuvo un valor de 1.852 eigenvalue (autovalor), y la varianza explicada por las medidas fue del 52.23% tal y como se recomienda que sea.¹⁴¹

Por otra parte, el valor del estadístico Yen-Q3 mostró que no hubo dependencia local en los 14 ítems del SQVD. ya que ninguno de los ítems obtuvo una correlación residual mayor de 0.2 por encima de la correlación promedio.¹⁶⁶ Por lo tanto esto indica que los 14 ítems del cuestionario deben estar presentes en el mismo.

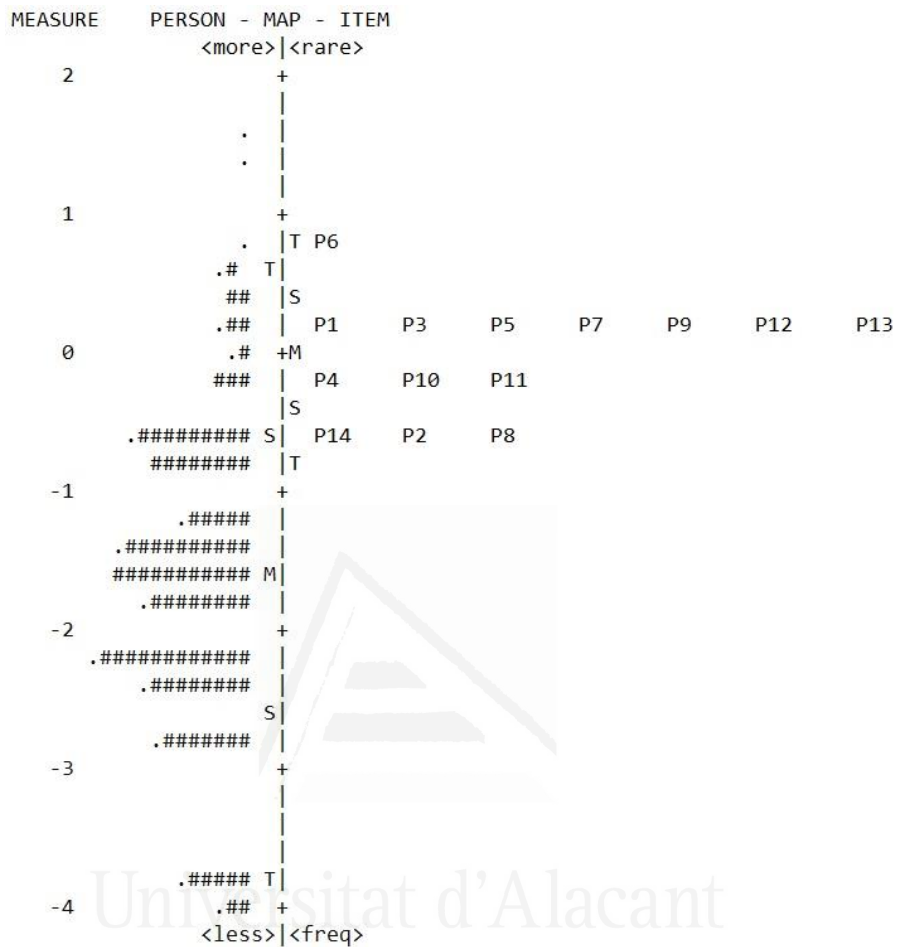


Figura 10. Mapa personas-ítem para el SQVD. Los sujetos están representados a la izquierda de la línea discontinua con el símbolo "#" (representa a 3 sujetos) y "." (representa entre 1 y 2 sujetos). A la derecha de la línea discontinua se muestran los ítems de la versión SQVD de 14 ítems con su número (P y Número del ítem). M indica el valor medio (a la izquierda la habilidad de las personas y a la derecha la dificultad de los ítems). S muestra una desviación estándar de la media y T denota dos desviaciones estándar. En la parte superior de la figura se representa una mayor habilidad para las personas (mayor frecuencia de síntomas) y una mayor dificultad para los ítems.

1.5.3.7. Tabla de transformación de las puntuaciones

Utilizando el análisis Rasch también se obtuvo la tabla de transformación de datos ordinales a una medida de intervalo.¹⁴¹ Los resultados de la tabla de transformación se muestran en la Tabla 18, donde se exponen la escala ordinal y las puntuaciones correspondientes a nivel de intervalo.

La conversión se utilizaría de la siguiente manera: por ejemplo, un paciente con una puntuación ordinal original de 15 puntos, tendrá la puntuación de intervalo de 14.49 utilizando el mismo rango de escala. Al aplicar esta conversión, se puede informar de cambios en la variable de medida, puesto que de este modo se permite utilizar análisis con estadística paramétrica. Por tanto, con esta tabla de transformación, los usuarios pueden aumentar la precisión del SQVD. Para utilizar esta tabla de conversión, además de que el paciente debe completar los datos de todos los ítems, es decir, no deben faltar datos, la población de estudio debe ser similar a la utilizada en esta investigación. Este hecho se observa si se comparan la Tabla 12 (tabla de transformación obtenida en la segunda investigación con 125 sujetos) y la Tabla 18 (tabla de transformación obtenida con 306 sujetos), apreciándose pequeñas variaciones en las puntuaciones de intervalo debido a que las dos muestras no tienen exactamente las mismas características.

Medida en escala ordinal	Medida de intervalo	
Puntuación cruda	Logit	Escala
0	-4.91	0.00
1	-3.66	3.55
2	-2.91	5.70
3	-2.44	7.04
4	-2.09	8.05
5	-1.80	8.88
6	-1.54	9.60
7	-1.31	10.25
8	-1.10	10.85
9	-0.90	11.42
10	-0.71	11.96
11	-0.53	12.48
12	-0.35	12.99
13	-0.18	13.50
14	0.00	13.99
15	0.17	14.49
16	0.35	15.00
17	0.53	15.51
18	0.71	16.03
19	0.90	16.57
20	1.10	17.14
21	1.31	17.74
22	1.54	18.39
23	1.80	19.11
24	2.09	19.95
25	2.44	20.96
26	2.91	22.30
27	3.67	24.45
28	4.91	28.00

Tabla 18. Conversión de la puntuación del SQVD (306 pacientes) sin procesar (0 - 28) a una escala de intervalo en unidades logit y utilizando las métricas de la escala original.

1.5.3.8. Exactitud diagnóstica mediante curvas ROC

Para el análisis de las curvas ROC, a los 306 pacientes de la muestra se les dividió en sujetos sintomáticos y asintomáticos, utilizando como Gold-estándar los síntomas referidos por los pacientes durante la recogida de datos en la historia clínica. Para ello, se consideró que un sujeto era sintomático cuando éste refería al menos un síntoma durante la anamnesis. De este modo, el análisis de la exactitud diagnóstica se basó en la puntuación bruta del SQVD para ambos grupos.

La Figura 11 muestra la curva ROC obtenida en este estudio. El área bajo la curva obtenida tenía un valor de 0.836 ($p < 0.001$), con unos intervalos de confianza al 95% entre 0.792 y 0.879. Este resultado implica que el SQVD tiene una buena exactitud diagnóstica. De este modo, al evaluar los diferentes puntos de corte encontrados, se observó que el equilibrio entre la sensibilidad y especificidad se estableció para un punto de corte ≥ 6 , para el que los valores de sensibilidad y especificidad fueron de 0.759 y 0.783 respectivamente. Este punto de corte lo que significa es que si un paciente obtiene una puntuación ≥ 6 en el SQVD, esto indica que el sujeto presenta síntomas relacionados con una disfunción visual. Por lo tanto, éste deberá ser el punto de corte a utilizar en este cuestionario cuando se quiera conocer si el paciente presenta o no síntomas por cualquier anomalía visual.

1.5.3.9. Validez

En esta investigación, el estudio de la validez del SQVD se realizó evaluando la validez concurrente y la validez de grupos conocidos. La validez concurrente representa el nivel de correlación entre la puntuación del cuestionario y la puntuación de las medidas clínicas. En este estudio, se correlacionó la puntuación del SQVD con la

agudeza visual (AV) LogMAR mediante el coeficiente de correlación Rho de Spearman. Los resultados mostraron un coeficiente de correlación significativo de 0.246 ($p < 0.001$) entre la agudeza visual LogMAR y la puntuación del cuestionario. A pesar de existir una correlación significativa, la baja correlación encontrada implica que estas anomalías visuales no siempre están relacionadas con un problema de agudeza visual. Una disfunción refractiva sí puede estar asociada a un problema de agudeza visual. Pero, por ejemplo, aunque las disfunciones acomodativas y binoculares pueden provocar síntomas, no tiene por qué verse afectada la agudeza visual y estar disminuida, lo que confirma los resultados de esta baja correlación.

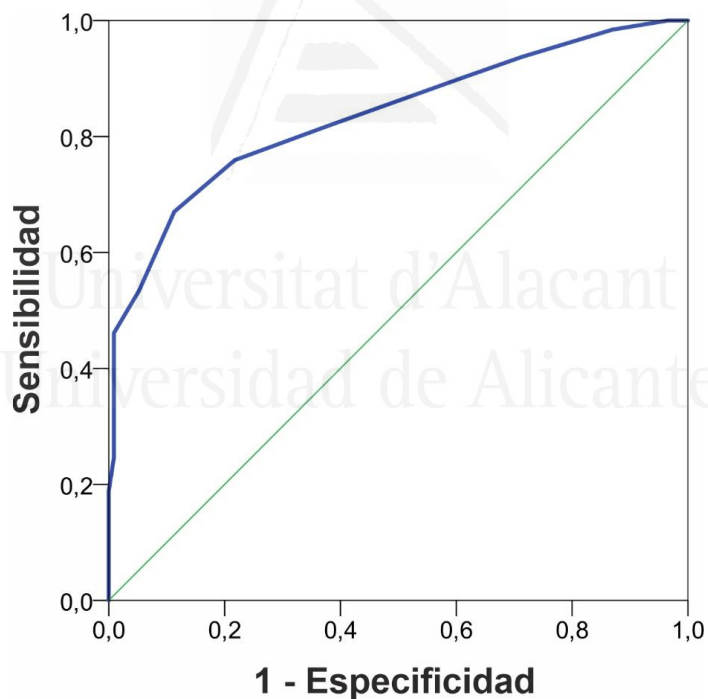


Figura 11. Curva ROC para el cuestionario SQVD. La curva ROC resume los resultados del análisis de sensibilidad y especificidad para varios puntos de corte. Cuanto mayor sea el área bajo la curva, mejor será la precisión del SQVD para detectar síntomas.

En cuanto a la validez de grupos conocidos, que es el grado en que el instrumento puede discriminar entre grupos clínicamente diferentes, se evaluó analizando las diferencias estadísticamente significativas encontradas entre las puntuaciones de dos grupos de pacientes (los que presentaban disfunciones visuales y los que no presentaban ninguna disfunción visual). Para ello, se utilizó la prueba U de Mann-Whitney. La puntuación media de SQVD para sujetos con cualquier tipo de disfunción visual fue de 8.41 ± 4.25 y de 3.82 ± 2.19 para sujetos sin disfunciones, mostrando diferencias estadísticamente significativas ($p < 0.001$) entre las puntuaciones de ambos grupos. Analizándolo por disfunciones, los sujetos con disfunciones refractivas obtuvieron una puntuación de 8.61 ± 4.36 y los sujetos con anomalías acomodativas y binoculares obtuvieron un valor medio de 8.06 ± 4.07 . Ambos grupos también presentaron diferencias estadísticamente significativas con el grupo de pacientes sin anomalías visuales ($p < 0.001$). Por tanto, según la validez de grupos conocidos, el instrumento muestra que es capaz de diferenciar sujetos con y sin disfunciones en función de su sintomatología.

Por lo tanto, los resultados de los dos tipos de validez evaluados en este estudio (validez concurrente y la validez de grupos conocidos), implican que el SQVD tiene una buena validez.

1.5.3.10. Repetibilidad

Para evaluar la repetibilidad (test-retest) del SQVD, se administró el cuestionario una semana después de la primera cumplimentación¹⁷¹ a una muestra de 50 pacientes elegidos aleatoriamente. Dicha muestra se calculó siguiendo los criterios establecidos por Walter et al.¹⁹³ para calcular el tamaño muestral en estudios de repetibilidad. Así, considerando un ICC esperado de 0.85, dos realizaciones del cuestionario, un ICC

mínimo aceptable de 0.70, un nivel de significación de 0.05 y una potencia del 80%, el tamaño mínimo de la muestra necesario resultó ser de 42 pacientes.

La repetibilidad test-retest mostró valores para el ICC de 0.857 ($p < 0.001$) con unos intervalos de confianza al 95% entre 0.710 y 0.933. Por lo tanto, el SQVD muestra que tiene muy buena repetibilidad. Este resultado es importante por ejemplo a la hora de evaluar cambios en los síntomas al considerar un tratamiento específico. El hecho de saber que el cuestionario es repetible permite que, si se utiliza en otro momento para monitorizar por ejemplo un tratamiento, los clínicos tengan la seguridad de que los resultados son fiables.

Todos estos resultados obtenidos en esta última parte de la tesis doctoral permiten comprender la principal contribución de toda esta investigación, ya que el cuestionario SQVD ofrece a los profesionales de la salud visual una herramienta para utilizar en la práctica clínica habitual a la hora de manejar las disfunciones visuales. No existen otros cuestionarios similares que puedan compararse con el SQVD. Existe un cuestionario desarrollado para niños en edad escolar con anomalías binoculares no estrábicas,¹²⁴ pero está diseñado para cuantificar su calidad de vida. Aunque tiene varios ítems con preguntas sobre varios síntomas visuales, la herramienta ha sido desarrollada para medir la calidad de vida en esta población infantil, siendo difícil la comparación con el SQVD. El cuestionario de Conlon¹³⁰ se desarrolló para medir el disconfort visual, pero la descripción de disconfort visual no está relacionada ni es comparable con los síntomas asociados a las disfunciones visuales, como ya se ha comentado anteriormente. Los ítems del cuestionario de Conlon tienen varios síntomas en la misma pregunta, lo que dificulta la comparación con otros instrumentos. En cualquier caso, varias preguntas tienen similitudes con diferentes ítems del SQVD, como síntomas relacionados con dolor de cabeza, visión borrosa, ojos irritados y problemas de lectura. Lo mismo ocurre al comparar el SQVD con el cuestionario CISS.^{131,132} Los síntomas

relacionados con dolor de cabeza, visión borrosa, ojos rojos, releer el texto y dar sueño se encuentran en ambos cuestionarios. Esto debería sugerir que el CISS podría usarse para evaluar síntomas en cualquier anomalía visual. Sin embargo, solo se ha desarrollado para insuficiencia de convergencia. Y aunque actualmente se ha realizado una adaptación transcultural para el idioma español utilizando el método Rasch,¹⁸⁶ el cuestionario original^{131,132} no fue validado con este método. Varios autores han utilizado el CISS para otras disfunciones visuales,^{62,178-180} pero esta manera de utilizarlo no es adecuada ya que solo se puede utilizar un cuestionario para la condición que se desarrolló y validó.^{181,182} De hecho, otros estudios han descrito determinados problemas y dificultades al utilizar el CISS en otras condiciones visuales.¹⁸³⁻¹⁸⁵

Todos estos problemas con los cuestionarios existentes corroboran la necesidad de contar con un cuestionario específicamente desarrollado y validado para cualquier disfunción visual, refractiva, acomodativa o binocular, como es el SQVD.

En cualquier caso, al igual que los estudios anteriores a éste, esta tercera parte de la tesis doctoral este estudio no está exenta de limitaciones. Por un lado, una de las limitaciones está relacionada con el targeting y la fiabilidad de los ítems. Ambos resultados implican que sería interesante analizar muestras más grandes de pacientes con niveles más altos de síntomas. De todos modos, el estudio piloto en el que se basó esta investigación¹⁹⁴ analizó una muestra de 125 sujetos y al utilizar en esta parte de la investigación una mayor muestra clínica con 306 sujetos, se demuestra que los valores del targeting y de la fiabilidad de los ítems mejoran, por lo que existe esa mejora.

Otra limitación del estudio estaría relacionada con la validez convergente. En esta investigación no se ha determinado este tipo de validez al no existir un cuestionario similar al SQVD que analice la sintomatología de todas las disfunciones visuales. Sin embargo, en futuras investigaciones sería interesante probar la validez discriminante

para comparar el SQVD con otros instrumentos relacionados con otras condiciones visuales.

A pesar de estas limitaciones, la fortaleza de este estudio es el reclutamiento de pacientes de una población clínica. A efectos clínicos, se podrá utilizar este instrumento para detectar la presencia y frecuencia de síntomas relacionados con cualquier tipo de disfunción visual en una muestra de similares características a la que se ha utilizado en esta investigación. Por todo ello, el principal aporte de este estudio es que este cuestionario será una ayuda para la detección de síntomas en la práctica clínica. El SQVD sería útil no solo para fines diagnósticos, ya que podría detectar con precisión la presencia de síntomas, sino también para monitorizar el tratamiento de las anomalías visuales. Así, los profesionales de la salud visual podrían detectar la mejoría de los pacientes mediante la desaparición de los síntomas o la disminución de su frecuencia.

En conclusión, los resultados de esta tercera investigación muestran que el SQVD tiene buenas propiedades psicométricas, buena exactitud diagnóstica, siendo un cuestionario válido y repetible para detectar la presencia y frecuencia de síntomas visuales relacionados con cualquier tipo de disfunciones visuales, disfunciones refractivas, acomodativas y binoculares. Por lo tanto, esta herramienta puede usarse en la práctica clínica para identificar sujetos con síntomas asociados con anomalías visuales y también es válida para su uso en estudios de investigación.

SECCIÓN 2:

TRABAJOS PUBLICADOS

Universitat d'Alicante
Universidad de Alicante

2.1. CAPÍTULO I: DELPHI METHODOLOGY FOR SYMPTOMATOLOGY ASSOCIATED WITH VISUAL DYSFUNCTIONS

Referencia: Cantó-Cerdán M, Cacho-Martínez P, García-Muñoz A. Delphi methodology for symptomatology associated with visual dysfunctions. Sci Rep. 2020;10:19403.

Autores:

1. Mario Cantó-Cerdán, MSc

Departamento de Óptica, Farmacología y Anatomía. Universidad de Alicante.

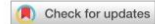
2. Pilar Cacho-Martínez, PhD

Departamento de Óptica, Farmacología y Anatomía. Universidad de Alicante.

3. Ángel García-Muñoz, PhD

Departamento de Óptica, Farmacología y Anatomía. Universidad de Alicante.

scientific reports



OPEN Delphi methodology for symptomatology associated with visual dysfunctions

Mario Cantó-Cerdán, Pilar Cacho-Martínez[✉] & Ángel García-Muñoz

To analyse what eyecare clinicians think about which symptoms are associated with refractive, accommodative and binocular dysfunctions, and which of them should be used in a questionnaire of visual symptomatology. A Delphi method was developed, using a coordinating group and a group of experts, and the process was conducted in three rounds. In the first round we compiled a list of 34 symptoms from the scientific literature and additional 10 suggested by the experts. These symptoms were categorized by each expert to the associated visual anomalies and working distance. In the second round, the relationship between each symptom and visual dysfunctions was analysed using a numeric scale. In the third round, the appearance or absence of the 44 symptoms in a questionnaire was assessed. Symptoms most frequently assigned by the experts to visual anomalies were related to near vision. Symptoms of blurred vision, difficulty focusing from one distance to another and close one eye obtained the highest mean score for refractive, accommodative and binocular disorders respectively. The experts were in agreement for 15 symptoms and in disagreement for 5 symptoms that should appear in a questionnaire. Delphi method has been used to identify the symptoms related to visual dysfunctions according to eyecare professionals and has allowed to arrive at appropriate symptoms to be asked for in a visual symptomatology questionnaire.

The presence of any uncorrected visual dysfunction (refractive, accommodative or vergence anomaly) may lead to an increase in visual symptoms which could impact on patient's visual comfort during the task. As symptoms may worsen with increased visual task, intervention will be necessary. The appropriate treatment of these visual problems may minimize or eliminate these symptoms.

These visual dysfunctions are commonly found in patients who attend optometry clinical practice¹⁻³ although there is disparity according to the symptomatology they produce and which symptoms are considered for their diagnosis⁴. There are also differences in both the way of asking for symptoms and how the severity of the symptoms are categorised⁵. Sometimes clinicians ask for symptoms using questionnaires or by means of the clinical history. When using clinical history, there are also differences, as several practitioners analyse symptoms using patient's descriptions of their case histories and in other situations symptoms are analysed on the basis of questions posed by the person who conducts the examination. Even when questionnaires are used, differences are shown as they use different scales to graduate the severity of symptoms. In any case, these differences do not influence the prescription of an appropriate treatment for any case⁶.

Recently, a systematic review⁷ has shown the wide disparity of symptoms related to different visual anomalies (refractive, accommodative and binocular ones) in the scientific literature and the differences between authors in methods for collecting this information. Most common symptoms may include *blurred vision*, *diplopia*, *headache* or *visual fatigue*, being 34 categories of symptoms, most of which are associated with near vision (Table 1). García-Muñoz et al.³ also found 11 different questionnaires for analysing visual symptoms⁷⁻¹⁸. Some questionnaires use dichotomous responses options and others by using multiple choice up to five responses for item. Nevertheless, only three questionnaires are psychometrically validated^{8,9,14,15}. Two of them are only related to convergence insufficiency (Convergence Insufficiency Symptoms Survey (CISS) for children¹⁴ and adults¹⁵ and Convergence Insufficiency and Reading Study (CIRS) parent version⁸. The Conlon survey⁹ measures visual discomfort but it is not associated with any particular visual anomaly (refractive, accommodative or binocular).

Due to the lack of availability of a generic symptom questionnaire to assess visual dysfunctions, this study aimed to use an approach to detect symptoms associated with visual anomalies which could be included in a questionnaire that would serve this purpose of aiding anomaly classification and measuring symptoms severity. Thus, clinicians could use this survey as an aid for diagnosing purposes and to monitor their treatment. For that,

Departamento de Óptica, Farmacología y Anatomía, Universidad de Alicante, Apartado 99, 03080 Alicante, Spain. ✉email: cacho@ua.es

Symptoms
Asthenopia
Avoid near tasks
Be distracted
Blurred vision
Change reading distance
Close one eye
Difficulty focusing from one distance to another
Difficulty performing schoolwork
Diplopia
Dry or gritty eyes
Excessive blinking
Excessive sensitivity to light
Extraordinary reading or writing posture
Eye turn noticed
Feel sleepy
Get faint colours around words
Head or book tilt
Headache
Inability to estimate distance accurately
Lack of concentration
Loss of place when reading
Ocular pain
Pulling eyes
Reading problems
Red eye
Rubbing of eyes
School performance problems
Sore eyes
Stomach ache or nausea
Tearing
Visual discomfort
Visual fatigue
Watery eyes
Words appear to move or jump at near vision

Table 1. 34 symptoms categories shown in García-Muñoz et al.⁶ review.

the first step would be to know which symptoms the clinicians would consider to be related to visual anomalies. And the Delphi method may be an efficient way to collect this information.

The Delphi technique is a formal consensus method used extensively in health care^{19,20} considered as an efficient way to collect information from a group of experts¹⁹. It is defined as a "method of structuring a group communication process that is effective in allowing a group of individuals, as a whole, to deal with a complex problem"^{21–23}. It is a prospective structured communication technique, essentially to obtain qualitative information by taking into consideration the opinion of each member of the panel.

The Delphi method allow experts to communicate their opinions and knowledge about a complex problem in order to explore options and potentially reach consensus even if they are in geographically dispersed areas²⁴. In fact, it has been shown that Delphi method is an useful tool to reach consensus in an area of uncertainty or lack of empirical evidence²⁵. This method has several characteristics which make it different from other consensus methods¹⁹: response anonymity (members do not know each other, which allows them not to be inhibited in their responses); iteration (the process occur in rounds where interactions between group members is made via questionnaire rather face to face communication) and controlled individual feedback (showing the distribution of the group's response). In addition to that, there is a statistical analysis using summary measures of the full group response. There is no standard method to calculate the number of experts for the Delphi technique. A sample of about fifteen has been suggested although larger panels may be used²⁶.

Taking into account these considerations, for our research, the application of the Delphi method may be useful to assess the opinion of clinicians regarding which visual symptoms may be related to visual anomalies. It was chosen over other consensus techniques because its ability to allow all experts equal participation and influence. The scientific evidence⁵ shows that the available questionnaires have not been developed taking into account the experience of the optical and clinical set-up practitioners. And in any case have been developed using

formal consensus methods as the Delphi method. So that using a Delphi method to know which symptoms are considered by the experts to assemble a robust questionnaire related to visual disorders, may be an important aid based in scientific evidence.

Therefore, the aim of this study is to identify which symptoms optometrists consider to be highly associated with presence of refractive, accommodative or binocular vision disorder, and of these symptoms, which ones should be included in a visual symptom questionnaire for diagnostic or treatment use.

Methods

This research was approved by the ethical committee of the University of Alicante. All methods were carried out in accordance with relevant guidelines and regulations.

Participants. For this study, we developed the Delphi method by establishing the coordinating and the expert group. The coordinating group comprised the authors of this manuscript. All of them were experienced optometrists and experts in public health. They were responsible for the overall coordinating of the study logistics. The coordinating role embraces functions such as modifying the study protocol where necessary, recruitment of experts, preparing the questionnaires of each round, analysing their responses as well as, interpreting the results from the study.

The experts' group were responsible for giving their opinions in each round. We initially invited 17 Spanish optometrists to participate in the study. All of them were professionals of recognised prestige with more than 15 years of experience, and with expertise in the field of evaluation and treatment of different vision anomalies. Three of them rejected the invitation, so the final panel of experts consisted of 14 members. They were selected according to the main characteristics the scientific literature recommends²⁷ to consider an expert for a Delphi technique, that is, according to their knowledge (several of them had a higher degree) and their experience (all of them had also clinical experience in several areas of vision). Of them, 5 optometrists developed their clinical practice in ophthalmological clinics and the other 9 in optical establishments so that this allowed to have different point of views from experts who examine different type of patients from their practices. All experts attended all age groups and different areas of visual health in their clinical practices. In this way, the group of experts included in the study covered different population groups. They were individually contacted via email in order to preserve anonymity. All experts gave their acceptance by informed consent to participate in this research.

Design and procedure. We designed a specific database to collect the information obtained from the panel of experts. As experts were from different geographic areas, this method allowed to get different ideas from several optometrists from all over Spain. Due to anonymity, possible negative influences of the dominant members of the group or the inhibition of a participant were avoided.^{28,29}

The coordinating group formulated the research question involved to the panel of experts and our application of the Delphi method involved three rounds. For each round there was an information exchange between the coordinating group and each expert individually. After each of them, the coordinating group collected the judgments and opinions of the experts and analysed them. Based on the responses collected and the interpretation of the coordinating group, a new questionnaire of the proposed problem was presented for the following round. As the coordinating group sent an abstract of the results obtained at the end of each round, before starting the next round, the experts knew the opinions obtained from all of them so that they could restate their answers depending on the results obtained in each round. With this procedure, a feedback was created through the analysis of the coordinating group, which allowed the flow of information among the experts and facilitated the establishment of a common language in each of the rounds. In addition, this approach allowed the experts the possibility of reflecting or reconsidering their opinion through their own approaches or from other experts as they could see the new information in the following round.

This procedure was done for each round until the coordinating group informed the experts of the panel's conclusions. Each round lasted two weeks. If the participants did not respond at this time or did not complete the questionnaire, they were sent a reminder email, thus avoiding biases in the study. All data was performed using the SPSS 20.0 statistical package.

First round. In the first round, we compiled a list of 34 symptoms (Table 1) from a systematic review³ about symptoms and visual anomalies. In this round, we developed a closed response questionnaire and experts were requested to classify the 34 symptoms according to the anomalies (refractive, accommodative, or binocular) which they are most likely to be associated with. The experts had also to classify these symptoms according to whether they thought they could be associated with distant or near vision. This association was considered when the answers were designated by more or equal than 50% of experts. Furthermore, in this first round the experts suggested to add 10 symptoms to the initial 34. These symptoms were considered when at least one expert recommended one of them. The 10 additional symptoms included: *clumsy/stumble with objects; difficulty completing assignments on time; fixation difficulty; forgetful; inability to maintain demanding visual effort activities continuously; irritability with visual effort activities; mist/spiderweb; neck and or back pain; pain in the area of the eyebrows and squint to see better*. Once the results were received, they were synthesised, analysed and sent to the experts together with the necessary documents to start the second round.

Second round. In this round the experts answered a closed response questionnaire in which each of them had to quantify the relationship between the 44 categories of symptoms and each of the dysfunctions. They had to rate the relationship by a scale of 0–10, with 0 being the least related and 10 the most related. Once all the answers

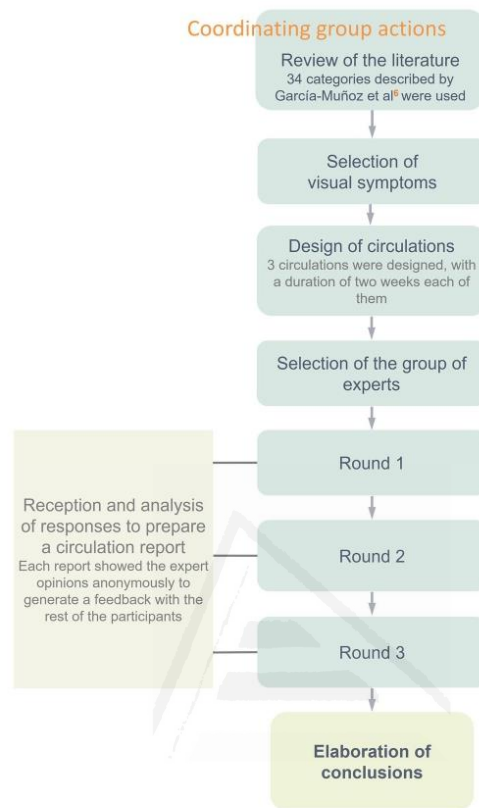


Figure 1. Representation of Delphi methodology used in the study.

were received, they were synthesised and analysed (by means of the mean and standard deviation), and results of this round were again sent to the experts together with the third circulation.

Third round. A closed answer questionnaire was also used. The experts had to indicate whether each of the 44 symptoms could appear in a visual symptom questionnaire. For this purpose, we used a Likert scale of 5 responses, where the experts had to answer their level of agreement to include each symptom in the questionnaire (*strongly disagree/disagree/neither agree or disagree/agree/strongly agree*). They had only to point out one of the specific responses for each symptom. When all responses were received, they were synthesised and analysed (by means of median and the interquartile range exposed through a box and whisker plot), and the third round results were sent back to the experts. Consensus for symptoms being in a questionnaire was assumed (by de coordination group) when the median was in the level of *strongly agree* in addition to have all scores between *neither agree or disagree* and *strongly agree*. Using this criterion, we can assure that no one expert disagrees with the inclusion of a symptom in the questionnaire. Consensus for symptoms not being in questionnaire was assumed when the median was in the level of *strongly disagree*, *disagree* and *neither agree or disagree*.

Afterwards, all data of the three rounds were analysed and a final report was revealed to the experts with the conclusions of the panel. Figure 1 shows the Delphi method developed by the coordinator group.

Results

Results of the first round were considered when the answers were designated by more or equal than 50% of experts. With these considerations, results show that 28 of the initial 34 symptoms (82.4%) were unique for refractive dysfunctions (Fig. 2); 23 of 34 symptoms (67.6%) were associated with accommodative dysfunctions (Fig. 3) and 28 of 34 (82.4%) were indicated as symptoms related to binocular dysfunctions (Fig. 4). Analysing how many symptoms overlap between categories, it is shown that 21 of 34 symptoms (61.8%) were related with

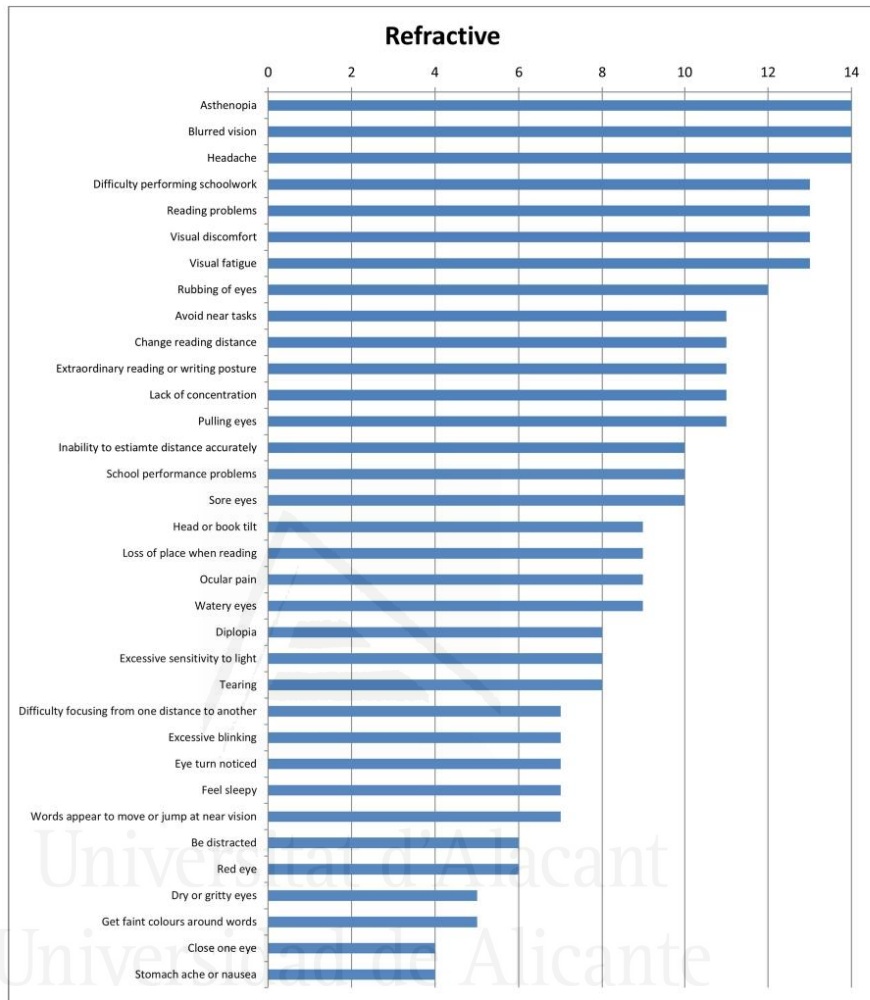


Figure 2. Results of round 1 for refractive dysfunctions. Number of experts who associate each of 34 initial symptoms with refractive dysfunctions.

the three categories of visual dysfunctions. However, only one symptom was associated with both refractive and accommodative anomalies. 5 of 34 symptoms (14.7%) were only related to both refractive and binocular dysfunctions. And there was not any symptom related to both accommodative and binocular disorders. First round results also show that experts related to distance vision 10 of 34 symptoms (Fig. 5) and most of symptoms (27 of 34 symptoms) were associated with near vision (Fig. 6).

Table 2 shows the results of the second round according to the scores given by the experts, associating each symptom with visual anomalies. Experts gave the high score to the symptom of *blurred vision* for refractive dysfunctions (9.64 ± 0.61), *difficulty focusing from one distance to another* for accommodative anomalies (9.43 ± 0.73) and *close one eye* for binocular ones (9.71 ± 0.59).

Figure 7 shows the results of the third round, in which the experts indicated their level of agreement about appearing each of the 44 symptoms in a questionnaire. Experts considered that 15 symptoms should be in a questionnaire, 5 symptoms should not appear and there was no consensus for 24 symptoms.

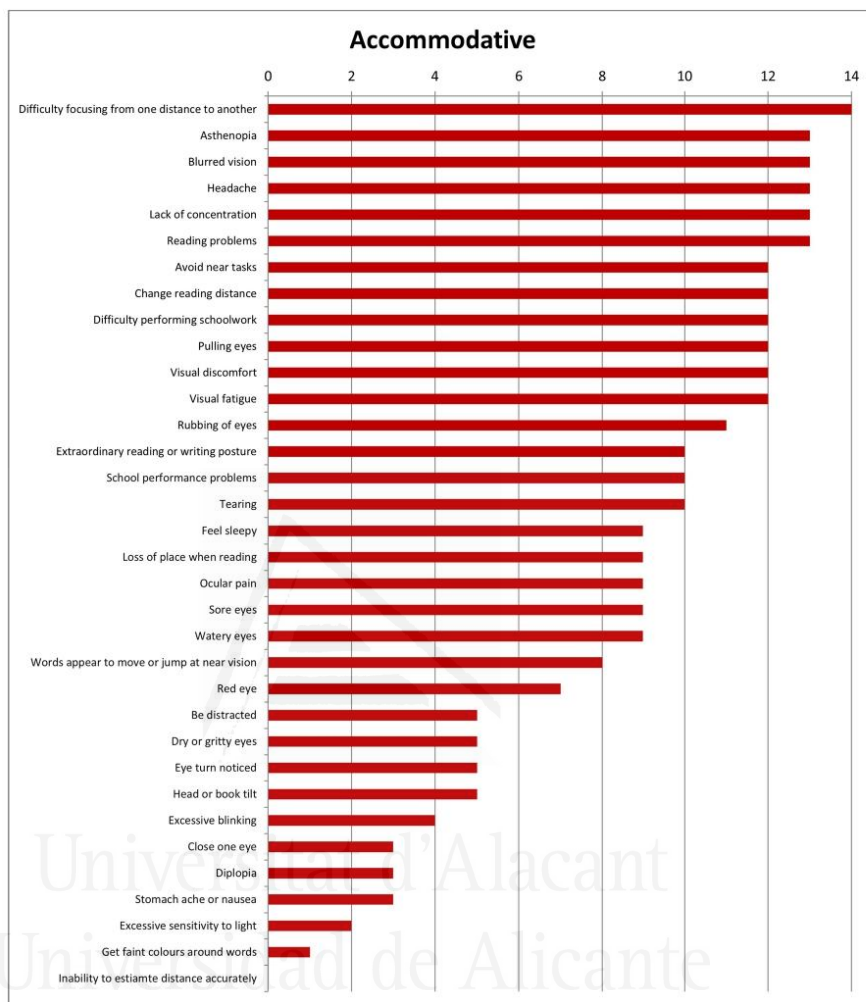


Figure 3. Results of round 1 for accommodative dysfunctions. Number of experts who associate each of 34 initial symptoms with accommodative dysfunctions.

Discussion

This research shows that Delphi method has been useful to identify the symptoms most likely to be associated with refractive, accommodative and binocular anomalies that should be included in a visual symptom questionnaire. The most associated symptoms with refractive anomalies were *asthenopia*, *blurred vision* and *headache*. For accommodative dysfunctions, *difficulty focusing from one distance to another* and for binocular anomalies, *diplopia* and *eye turn noticed*. Experts associated most of the symptoms to near vision. The panel of experts also gave the highest mean score to the symptom of *blurred vision* for refractive dysfunctions, *difficulty focusing from one distance to another* for accommodative anomalies and *close one eye* for binocular disorders. And they agreed with 15 symptoms should be in a questionnaire, 5 symptoms should not appear and there was no consensus for 24 symptoms.

First round results show that many symptoms related to accommodative dysfunctions are also associated with refractive anomalies, indicating that symptoms of both visual dysfunctions may be similar. This finding is similar to that observed in the study of Cacho-Martínez et al.⁴ where the authors found that uncorrected refractive error contaminates the symptoms of visual dysfunction, fundamentally when an accommodative dysfunction is present.

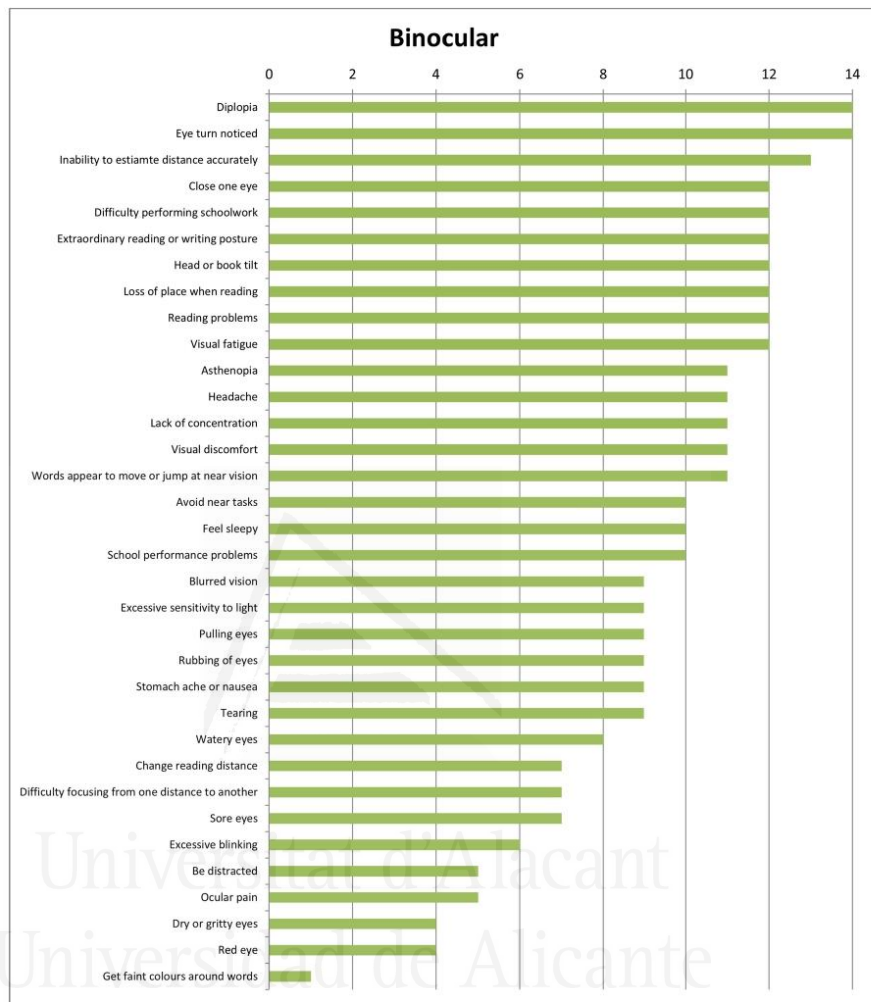


Figure 4. Results of round 1 for binocular dysfunctions. Number of experts who associate each of 34 initial symptoms with binocular dysfunctions.

On the other hand, the experts associated most of the symptoms to near vision, highlighting those of *diplopia*, *visual discomfort* and *visual fatigue*. These results are similar to those obtained in the systematic review of Garcia-Muñoz et al.⁵ in which the authors showed that most of symptoms are associated with near vision.

In the second round, when doing the analysis for dysfunctions and considering the score given by the experts, we can observe that *blurred vision* was the symptom that obtained the highest mean for refractive dysfunctions. However, it also received a high score for accommodative dysfunctions, showing that it would not be exclusive for refractive anomalies³⁰, being able to be related to a refractive or accommodative anomaly^{4,31–33}. This agrees with the fact that the symptom of *blurred vision* appears in most of visual questionnaires^{7–18} showed by scientific literature. We can also observe for refractive dysfunctions that the symptom of *squint to see better* is the other symptom which obtained a high score and the only one in which the experts agree that is specific for refractive dysfunctions. This is one of the ten symptoms that experts added in the first round. However, the scientific evidence has shown that it only appears in two questionnaires^{9,11}.

For accommodative anomalies, the highest mean was obtained by the symptom of *difficulty focusing from one distance to another* which, according to the experts, seems to be specific for these anomalies. However, it only appears in one questionnaire¹⁶, although this survey is related to accommodative dysfunctions. There was

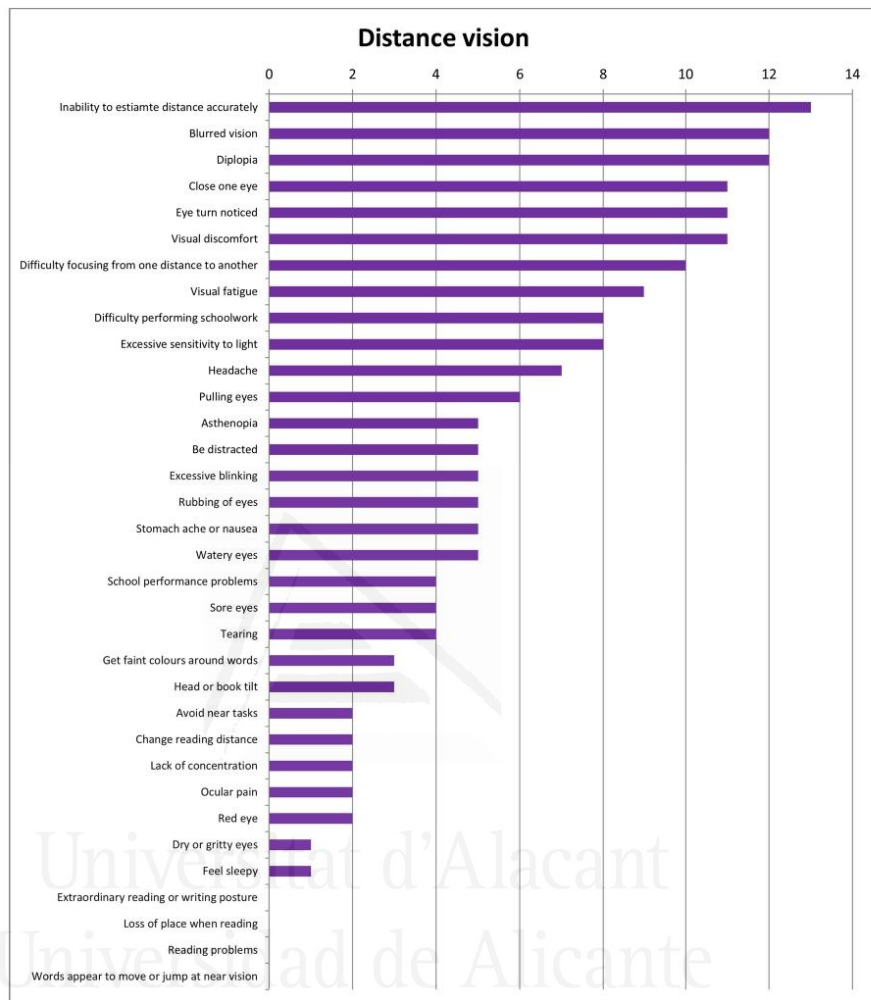


Figure 5. Results of round 1 for distance vision. Number of experts who associate each of 34 initial symptoms with distance vision.

another symptom (*asthenopia*) which got a high mean, not only for accommodative anomalies but also, for the other dysfunctions, showing that experts gave importance to it for all visual anomalies. However, this symptom does not appear in any of existing questionnaires⁷⁻¹⁸.

For binocular anomalies, the symptoms with the highest mean scores were *close one eye*, *diplopia* and *eye turn noticed*. Only one of the three symptoms (that is *diplopia*) appears in 8⁷⁻¹⁵ out of 11 existing questionnaires⁷⁻¹⁸. However, the symptom of *close one eye* only appears in the 19 Item College of Optometrist in Vision Development Quality of Life (COVD-QOL)¹⁷ while *eye turn noticed* does not appear in any questionnaire.

In general, the symptoms *dry or gritty eyes*, *forgetful* and *get faint colours around words* had low mean scores for all anomalies. This finding may be due to the low importance the experts assigned to those symptoms.

Furthermore, results from second round show that for binocular anomalies there were more symptoms (15 that is, 34%) with mean scores greater than 8. This result is higher than that found for refractive and accommodative dysfunctions which each one had 7 symptoms (15.91%) with scores greater than 8. These results indicate that the experts gave more importance to the symptoms associated with binocular dysfunctions than refractive and accommodative ones. This situation agrees with the results found by Cacho-Martínez et al.⁴ where for clinical population it was observed a greater association between having symptoms and presenting a binocular

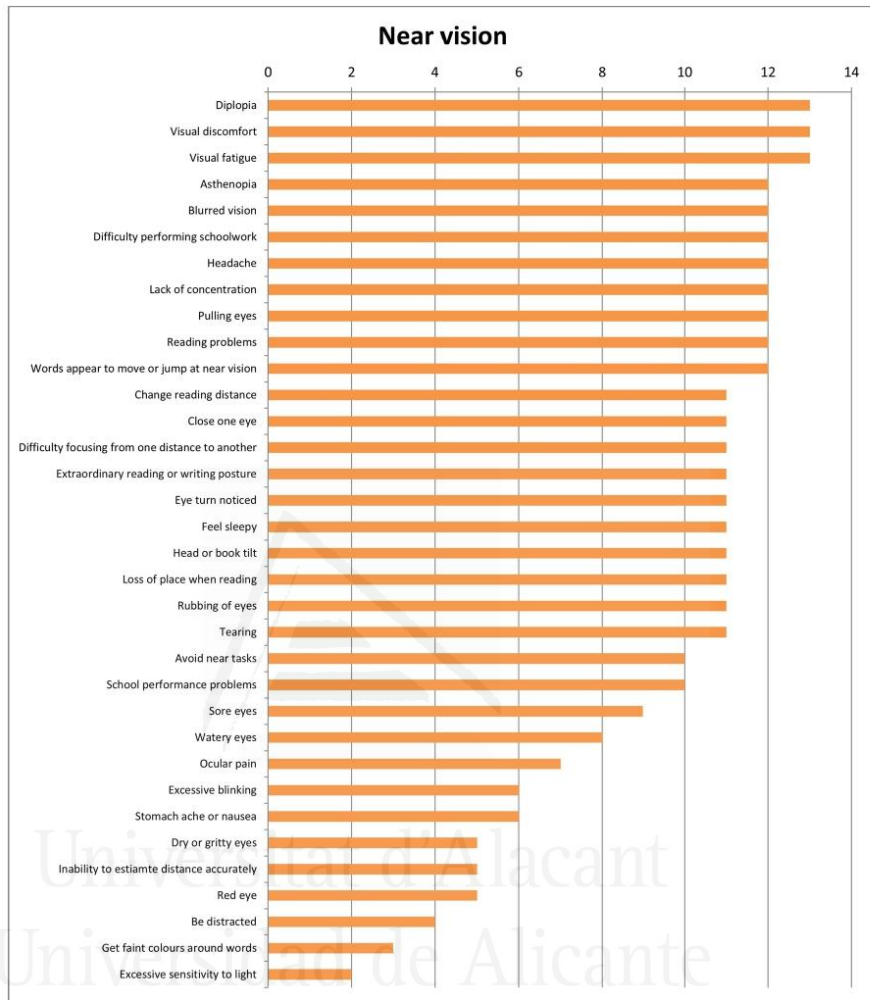


Figure 6. Results of round 1 for near vision. Number of experts who associate each of 34 initial symptoms with near vision.

dysfunction. In the same way, the scientific literature⁵ has also shown that most symptoms taken into account by the authors are related to binocular dysfunctions.

In any case, it is clear that the same symptom may be related to any anomaly. And although it seems that symptoms may be more associated with binocular dysfunctions, this does not imply that they are specific to a disorder. The fact that there are no specific symptoms for each visual dysfunction, suggests that a questionnaire may be developed for any visual dysfunction, and not a specific one for each visual anomaly.

In the third round, considering the established consensus, there was a great majority of agreement among experts that 15 of the initial 44 symptoms were important and should appear in a questionnaire. Figure 7 shows these symptoms. Of these 15 symptoms, 11 of them appear in most of the questionnaires published in the scientific literature⁷⁻¹⁷. However, the other four of them (*asthenopia*, *eye turn noticed*, *fixation difficulty* and *pain in the area of eyebrows*) do not appear in any survey. This finding shows the utility of the Delphi process, as it can be characterised symptoms commonly reported in clinical practice but not considered by the existing literature.

The experts also agreed that five symptoms were not important and should not be on the questionnaire. Of these five, two of them were included by the panel of experts (*forgetful* and *mist/spiderweb*), appearing only the symptom of *forgetful* in a questionnaire, particularly in the 19 item COVID-QOL. The remaining three symptoms

	Refractive		Accommodative		Binocular	
	Mean	SD	Mean	SD	Mean	SD
Asthenopia	8.93	1.33	9.14	1.36	8.71	1.71
Avoid near tasks	7.36	2.06	7.79	1.82	7.57	2.06
Be distracted	4.93	3.13	4.57	3.92	4.86	3.68
Blurred vision	9.64	0.61	7.79	2.14	4.93	2.55
Change reading distance	8.21	1.37	8.07	2.52	6.93	2.52
Close one eye	4.29	3.30	3.71	2.68	9.71	0.59
Clumsy/ Stumble with objects	5.64	2.89	2.50	2.90	7.64	2.87
Difficulty completing assignments on time	6.93	1.71	7.64	2.02	7.64	1.95
Difficulty focusing from one distance to another	6.64	2.69	9.43	0.73	5.64	3.15
Difficulty performing schoolwork	8.14	1.41	8.21	1.70	8.21	1.61
Diplopia	4.00	2.88	3.50	3.68	9.14	1.06
Dry or gritty eyes	4.21	2.57	3.79	2.83	2.86	2.82
Excessive blinking	5.71	2.08	5.79	2.24	6.71	2.91
Excessive sensitivity to light	5.50	3.31	2.93	2.63	7.36	3.33
Extraordinary reading or writing posture	6.50	2.56	7.36	2.35	8.21	2.04
Eye turn noticed	4.00	3.40	3.50	2.99	9.57	0.73
Feel sleepy	5.50	3.35	6.14	3.48	5.43	3.22
Fixation difficulty	5.57	3.22	7.00	1.85	8.07	1.33
Forgetful	1.43	1.72	1.79	2.34	2.21	2.93
Get faint colours around words	4.29	3.30	3.21	3.41	3.07	3.24
Head or book tilt	5.00	2.27	4.00	2.45	8.36	1.76
Headache	7.92	3.22	7.79	2.04	8.21	1.78
Inability to estimate distance accurately	5.57	2.61	3.14	2.64	8.79	1.82
Inability to maintain demanding visual effort activities continuously	7.79	1.47	8.43	1.35	8.50	1.35
Irritability with visual effort activities	6.36	2.50	7.36	1.87	7.57	2.66
Lack of concentration	7.29	2.19	7.79	1.66	8.00	1.56
Loss of place when reading	5.50	2.97	6.07	2.68	8.21	1.82
Mist/Spiderweb	7.29	1.67	5.14	2.82	2.21	2.86
Neck and/or back pain	3.64	3.08	3.50	2.87	5.57	2.87
Ocular pain	5.36	2.35	5.93	2.74	4.93	2.89
Pain in the area of the eyebrows	6.00	2.20	5.57	3.06	4.86	3.00
Pulling eyes	6.64	2.19	7.07	2.52	6.00	2.88
Reading problems	6.93	2.52	7.93	2.09	8.07	2.09
Red eye	5.36	2.22	5.14	2.67	4.00	2.88
Rubbing of eyes	7.36	1.67	7.21	2.57	7.07	2.68
School performance problems	5.79	3.71	6.21	3.78	6.36	3.64
Sore eyes	5.71	2.68	5.43	2.90	4.93	3.39
Squint to see better	9.36	0.97	6.36	3.24	3.21	3.07
Stomach ache or nausea	4.29	2.37	4.57	3.20	6.86	2.87
Tearing	6.14	2.36	6.29	2.43	6.36	2.64
Visual discomfort	8.14	1.41	8.29	1.79	8.43	1.40
Visual fatigue	8.14	1.77	8.29	2.58	7.93	2.49
Watery eyes	5.93	2.43	5.64	2.72	5.71	3.49
Words appear to move or jump at near vision	6.71	2.05	7.29	2.52	6.64	3.04

Table 2. Results of round 2. Mean and standard deviation (SD) of scores given by the panel of experts, relating each of the 44 symptoms to visual dysfunctions (refractive, accommodative and binocular anomalies). Score scale from 0 to 10.

(*be distracted*, *get faint colours around words* and *stomach ache or nausea*) appear in several questionnaires. The symptom of *stomach ache or nausea* is included in the CISS symptom questionnaire¹³. The symptom of *get faint colours around words* is incorporated in the Adler questionnaire¹¹ and *be distracted* appears in the 19 item COVID-QOL¹⁷. However, the panel of experts considered that although these four symptoms are used in the scientific literature, they have little importance in the clinical setting and should therefore be omitted. This finding shows that symptoms considered by existing literature are not reported by the experts when a Delphi process is used.

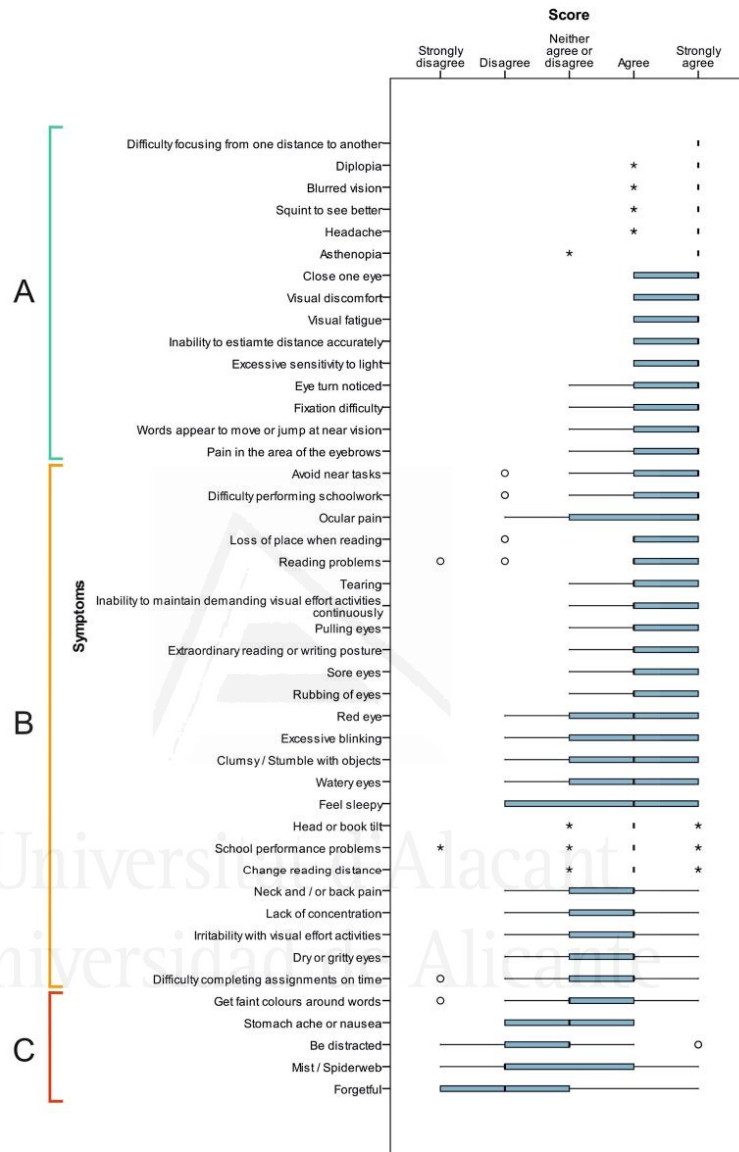


Figure 7. Results of round 3. Box and whisker plot showing the degree of agreement of experts regarding each of 44 symptoms to be in a visual symptom questionnaire. The width of the box shows the interquartile range (IQR). The black line in the box marks the median of the responses (which is the 50th percentile). The left line of the box shows the lower quartile (25th percentile) and the right line of the box represents the upper quartile (75th percentile). The left whisker indicates the lowest value obtained (a lower value within 1.5 IQR) and the right whisker shows the highest value obtained (which is the upper value within 1.5 IQR). The asterisk represents an outlier which is 1.5–3 IQR and the circle shows an outlier > 3 IQR. (A) 15 symptoms for which experts had an agreement for being in a questionnaire. (B) 24 symptoms with discrepancies for being in a questionnaire. (C) 5 symptoms for which experts had an agreement for not being in a questionnaire.

For the remaining 24 symptoms, the experts had discrepancies regarding their inclusion in a questionnaire, so they should be carefully analysed for example for a future questionnaire. For this reason, those symptoms with discrepancies could be initially considered when developing a future questionnaire. Then, the appropriate psychometric methods would aid to decide to include them or not.

Of these 24 symptoms there also were discrepancies between experts who practiced in an optometric clinic and those of optical centres. Experts from optical centers, highlighted that the following 10 symptoms had a great importance when asking for dysfunctions: *clumsy/ stumble with objects, difficulty completing assignments on time, dry or gritty eyes, excessive blinking, feel sleepy, neck and/or back pain, ocular pain, red eye, sore eyes, and watery eyes*. However, experts who practiced in optometric clinics did not consider them important. This lack of agreement must be due to the different type of patients attended in each centre, but it does not have to affect to consider or not those symptoms.

When reviewing the three rounds, we can see that there is little variation in the importance that experts give to the symptoms. The 15 symptoms which experts agreed with appearing in a questionnaire, all of them were considered important in the first or second round. In addition, something similar happens with the five symptoms that the experts considered should not appear in a questionnaire. That is, symptoms as *be distracted, get faint colours around words and stomach ache or nausea* had low scores in the first and second round for all dysfunctions (although *stomach ache or nausea* obtained better score for binocular anomalies). The other two symptoms, *mist/spiderweb* and *forgetful* were not in the first round and were added by the experts in the second round. In the case of *forgetful*, the experts agreed with not appearing in a questionnaire and in fact, was not important in the first and second round. However, a particular situation occurs for the symptom of *mist/spiderweb* for which the experts considered not important to be in a questionnaire and gave low scores for accommodative and binocular disorders but an acceptable score for refractive anomalies.

The minor variation between rounds shows that from the beginning, the experts knew which symptoms could be the most important. Certainly, there were minimal variations, usual situation when carrying out a Delphi method. In fact, the importance of this method is that experts can change impressions anonymously and reflect on their position, modifying it throughout the process. In any case, the research has shown that experts had clear their position on several symptoms, with discrepancies for other symptoms.

A limitation of the Delphi method is that experts may be influenced by the scoring of others which may bias the results. To avoid this, the identities of the experts were not shown. Another limitation has been the fact that optometrists representing the entire geographical area of the country have not been included in the final panel of experts. Although the choice of experts was initially trying to cover all the possible geographical areas, there were three experts who refused to participate. In any case the geographical area which they represented were only small areas.

Despite the outlined limitations, the strength of this study is the high response rate (100%) off the experts and their good participation, as all of them answered the three rounds of the Delphi methodology.

In conclusion, this Delphi study has been useful to identify the symptoms related to visual dysfunctions, both in patients from optometric clinics and optical centers, so it can be considered the starting point to develop a tool which may facilitate the collection and classification of these symptoms. This tool could be used for diagnosing and treatment purposes.

Received: 31 July 2020; Accepted: 27 October 2020

Published online: 10 November 2020

References

- Lara, F., Cacho, P., Garcia, A. & Megias, R. General binocular disorders: prevalence in a clinic population. *Ophthalm. Physiol. Opt.* **21**, 70–74 (2001).
- García-Muñoz, A., Carbonell-Bonete, S., Cantó-Cerdán, M. & Cacho-Martínez, P. Accommodative and binocular dysfunctions: prevalence in a randomised sample of university students. *Clin. Exp. Optom.* **99**, 313–321 (2016).
- Hashemi, H., Nabovati, P., Yekta, A., Shokrollahzadeh, F. & Khabazkhoob, M. The prevalence of refractive errors among adult rural populations in Iran. *Clin. Exp. Optom.* **101**, 84–89 (2018).
- Cacho-Martínez, P., Cantó-Cerdán, M., Carbonell-Bonete, S. & García-Muñoz, A. Characterization of visual symptomatology associated with refractive, accommodative, and binocular anomalies. *J. Ophthalmol.* [DOI:https://doi.org/10.1155/2015/895803](https://doi.org/10.1155/2015/895803) (2015).
- García-Muñoz, A., Carbonell-Bonete, S. & Cacho-Martínez, P. Symptomatology associated with accommodative and binocular vision anomalies. *J. Optom.* **7**, 178–192 (2014).
- Cacho-Martínez, P., García-Muñoz, A. & Ruiz-Cantero, M. T. Treatment of accommodative and nonstrabismic binocular dysfunctions: a systematic review. *Optometry* **80**, 702–716 (2009).
- Russell, G. E. & Wick, B. A prospective study of treatment of accommodative insufficiency. *Optom. Vis. Sci.* **70**, 131–135 (1993).
- Borsting, E., Rouse, M. & De Land, P. Prospective comparison of convergence insufficiency and normal binocular children on CIRS symptom surveys. Convergence Insufficiency and Reading Study (CIRS) group. *Optom. Vis. Sci.* **76**, 221–228 (1999).
- Conlon, E. G., Lovegrove, W. J., Chekaluk, E. & Pattison, P. E. Measuring visual discomfort. *Vis. Cognit.* **6**(6), 637–663 (1999).
- Birnbaum, M. H., Soden, R. & Cohen, A. H. Efficacy of vision therapy for convergence insufficiency in an adult male population. *J. Am. Optom. Assoc.* **70**, 225–232 (1999).
- Adler, P. Efficacy of treatment for convergence insufficiency using vision therapy. *Ophthalmic. Physiol. Opt.* **22**, 565–571 (2002).
- Gallaway, M., Scheiman, M. & Malhotra, K. The effectiveness of pencil pushups treatment for convergence insufficiency: a pilot study. *Optom. Vis. Sci.* **79**, 265–267 (2002).
- Borsting, E. *et al.* Association of symptoms and convergence and accommodative insufficiency in school-age children. *Optometry* **74**, 25–34 (2003).
- Borsting, E. *et al.* Validity and reliability of the revised convergence insufficiency symptom survey in children aged 9 to 18 years. *Optom. Vis. Sci.* **80**, 832–838 (2003).
- Rouse, M. *et al.* Validity and reliability of the revised convergence insufficiency symptom survey in adults. *Ophthalmic. Physiol. Opt.* **24**, 384–390 (2004).

16. Sterner, B., Gellerstedt, M. & Sjöström, A. Accommodation and the relationship to subjective symptoms with near work for young school children. *Ophthalmic. Physiol. Opt.* **26**, 148–155 (2006).
17. Vaughn, W., Maples, W. C. & Hoenes, R. The association between vision quality of life and academics as measured by the College of Optometrists in Vision Development Quality of Life questionnaire. *Optometry* **77**, 116–123 (2006).
18. Rouse, M. *et al.* Academic behaviors in children with convergence insufficiency with and without parent-reported ADHD. *Optom. Vis. Sci.* **86**, 1169–1177 (2009).
19. Dalkey, N. C. An experimental study of group opinion: the Delphi method. *Futures* **1**, 408–426 (1969).
20. Jones, J. & Hunter, D. Consensus methods for medical and health services research. *BMJ* **311**, 376–380 (1995).
21. Varela-Ruiz, M., Diaz-Bravo, L. & Garcia-Durán, R. Descripción y usos del método Delphi en investigaciones del área de la salud. *Inv. Ed. Med.* **1**(2), 90–95 (2012).
22. Meyrick, J. The Delphi method and health research. *Health Educ.* **103**, 7–16 (2003).
23. Dalkey, N. C. & Helmer, O. An experimental application of the delphi method to the user of experts. *Manage. Sci.* **9**, 458–477 (1963).
24. Rowe, J., Wright, G. & Bolger, F. Delphi: a reevaluation of research and theory. *Technol. Forecast. Soc. Chang.* **39**, 235–251 (1991).
25. Pill, J. The Delphi method: substance, context, a critique and an annotated bibliography. *Socioecon. Plan. Sci.* **5**(1), 57–71 (1971).
26. McMillan, S. S., King, M. & Tully, M. P. How to use the nominal group and Delphi techniques. *Int. J. Clin. Pharm.* **38**, 655–662 (2016).
27. Baker, J., Lovell, K. & Harris, N. How expert are the experts? An exploration of the concept of “expert” within Delphi panel techniques. *Nurse Res.* **14**, 59–70 (2006).
28. Keeney, S., Hasson, F. & McKenna, H. Consulting the oracle: ten lessons from using the Delphi technique in nursing research. *J. Adv. Nurs.* **53**, 205–212 (2006).
29. Powell, C. The Delphi technique: myths and realities. *J. Adv. Nurs.* **41**, 376–382 (2003).
30. Dwyer, P. & Wick, B. The influence of refractive correction upon disorders of vergence and accommodation. *Optom. Vis. Sci.* **72**, 224–232 (1995).
31. Cacho-Martínez, P., García-Muñoz, A. & Ruiz-Cantero, M. T. Do we really know the prevalence of accommodative and nonstrabismic binocular dysfunctions?. *J. Optom.* **3**, 185–197 (2010).
32. Chrousos, G. A., O'Neill, J. F., Lueth, B. D. & Parks, M. M. Accommodation deficiency in healthy young individuals. *J. Pediatr. Ophthalmol. Strabismus* **25**, 176–179 (1988).
33. Rutstein, R., Daum, K. & Amos, J. Accommodative spasm: a study of 17 cases. *J. Am. Optom. Assoc.* **59**, 527–538 (1988).

Acknowledgements

We wish to thank the participating experts for their collaboration in this methodology: Luis Bielsa-i-Elies; Natalia Cano-González; Enrique González-Díaz-Obregón; Vicente A. Fernández-Sánchez; Francisco Lara-Lacárcel; Margarita León-López; Rafael Pérez-Cambrodi; Mar Pérez-Crespo-Gómez; José Miguel Pelegrín-Sánchez; David P. Piñero-Llorens; Germán Rodríguez-Hernández; Vicente Rodríguez-Pascual; Francisco Valderas-Jiménez; Teyma Valero-Pérez.

Author contributions

M.C.C.: conceived and designed the study, analyzed and interpreted the data. P.C.M.: conceived and designed the study, analyzed and interpreted the data, and drafted the manuscript, which was revised by all the authors. A.G.M.: conceived and designed the study, analyzed and interpreted the data. All the authors approved its final version.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to P.C.-M.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020

2.2. CAPÍTULO II: RASCH ANALYSIS FOR DEVELOPMENT AND REDUCTION OF SYMPTOM QUESTIONNAIRE FOR VISUAL DYSFUNCTIONS (SQVD)

Referencia: Cantó-Cerdán M, Cacho-Martínez P, Lara-Lacárcel F, García-Muñoz A. Rasch analysis for development and reduction of Symptom Questionnaire for Visual Dysfunctions (SQVD). Sci Rep 2021; 11:14855.

Autores:

1. Mario Cantó-Cerdán, MSc.

Departamento de Óptica, Farmacología y Anatomía. Universidad de Alicante.

2. Pilar Cacho-Martínez, PhD.

Departamento de Óptica, Farmacología y Anatomía. Universidad de Alicante.

3. Francisco Lara-Lacárcel, PhD.

Departamento de Oftalmología, Optometría, Otorrinolaringología, y Anatomía Patológica. Universidad de Murcia.

4. Ángel García-Muñoz, PhD.

Departamento de Óptica, Farmacología y Anatomía. Universidad de Alicante.



OPEN

Rasch analysis for development and reduction of Symptom Questionnaire for Visual Dysfunctions (SQVD)

Mario Cantó-Cerdán¹, Pilar Cacho-Martínez^{1✉}, Francisco Lara-Lacárcel² & Ángel García-Muñoz¹

To develop the Symptom Questionnaire for Visual Dysfunctions (SQVD) and to perform a psychometric analysis using Rasch method to obtain an instrument which allows to detect the presence and frequency of visual symptoms related to any visual dysfunction. A pilot version of 33 items was carried out on a sample of 125 patients from an optometric clinic. Rasch model (using Andrich Rating Scale Model) was applied to investigate the category probability curves and Andrich thresholds, infit and outfit mean square, local dependency using Yen's Q3 statistic, Differential item functioning (DIF) for gender and presbyopia, person and item reliability, unidimensionality, targeting and ordinal to interval conversion table. Category probability curves suggested to collapse a response category. Rasch analysis reduced the questionnaire from 33 to 14 items. The final SQVD showed that 14 items fit to the model without local dependency and no significant DIF for gender and presbyopia. Person reliability was satisfactory (0.81). The first contrast of the residual was 1.908 eigenvalue, showing unidimensionality and targeting was -1.59 logits. In general, the SQVD is a well-structured tool which shows that data adequately fit the Rasch model, with adequate psychometric properties, making it a reliable and valid instrument to measure visual symptoms.

The presence of any type of visual dysfunction, refractive, accommodative and/or binocular, may cause the appearance of visual symptoms that affect the level of comfort of a patient when performing a visual task¹.

The scientific literature shows disparity regarding the symptoms associated with the diagnosis of these visual anomalies, the way for asking the patient about them and how to classify the degree of severity^{2,3}. In a scoping review published by García-Muñoz et al.³ it was shown that there are differences when compiling information and classifying symptoms of the anomalies. In that review it was shown that the most common symptoms related to these anomalies may include diplopia, movement or blinking of the words in near vision, headaches, visual fatigue and blurred vision and that these symptoms could be classified into 34 different categories, all of them fundamentally related to near vision. In this review it was also observed that there are a total of 11 questionnaires in the scientific literature⁴⁻¹⁵. Some use dichotomous questions and others multiple choice with up to 5 responses per item. Of the 11 questionnaires, only three of them were validated^{5,10-12}. One of them (the Conlon survey) referring to visual discomfort not associated to any specific dysfunction⁵ and the CISS V-15 and its version for parents (CIRS parent version) to convergence insufficiency¹⁰⁻¹².

Scientific literature has also shown specific validated questionnaires to collect the different symptoms which may be associated to computer screens or video terminals^{16,17}, although this type of questionnaires does not cover all situations of everyday life since they are specific to these devices.

As it can be seen, no symptom questionnaire is currently available asking questions which could be related to any type of visual dysfunction (refractive, accommodative or binocular) which may provoke symptoms. The available questionnaires have been developed for specific visual anomalies such as convergence insufficiency¹¹, have been related to visual discomfort in general⁵ or have been focused on particular samples as computer users^{16,17} and only some of them have been psychometric validated using Rasch analysis^{5,16,17}. However, we consider it should be interesting to have a questionnaire which would allow the detection of any visual symptoms associated with any type of visual dysfunction. The instrument would broadly cover all cause of these symptoms, including

¹Departamento de Óptica, Farmacología y Anatomía, Universidad de Alicante, Alicante, Spain. ²Departamento de Oftalmología, Optometría, Otorrinolaringología y Anatomía Patológica, Universidad de Murcia, Murcia, Spain. ✉email: cacho@ua.es

refractive, accommodative and binocular. That is, a questionnaire which could detect the presence or absence of visual symptoms and could measure their frequency. This tool would be useful for clinicians as they could accurately detect the presence of symptoms, supporting the diagnosing process. Furthermore, this questionnaire could be used by professionals to calibrate the visual symptomatology when monitoring the treatment of visual anomalies. Clinicians would be able to detect the improvement of patients by means of symptoms' disappearance or through the decrease of their frequency.

In this sense, we are developing a questionnaire that covers this need. At the beginning of the research, we developed a pilot questionnaire using Delphi methodology in which various experts discussed the inclusion of different categories of visual symptoms in a questionnaire¹⁵. This pilot questionnaire on visual symptomatology is the starting point for this research and for its development, a psychometrical evaluation using Rasch analysis is needed. There are several statistical methods to psychometrically analyze an instrument: The Classical Test Theory (CTT) or the Item Response Theory (IRT). CTT has several limitations which have given place to use the IRT. IRT refers to probabilistic measurement models according to the number of parameters analyzed so that Rasch analysis is used when 1-parameter is studied, that is, when unique construct exists. The strength of Rasch methodology is that it allows conjoint measurement of persons and items on the same dimension or construct¹⁹. This method is recommended in this type of instruments²⁰, as it provides an idea of the internal consistency of the scale and can match the item difficulty with the person's ability²¹.

Therefore, the aim of this study is to develop the Symptom Questionnaire for Visual Dysfunctions (SQVD) and to perform a psychometric analysis using Rasch method to obtain an instrument which allows to detect the presence and frequency of visual symptoms related to any visual dysfunction.

Methods

Design of the pilot questionnaire. In this study the Symptom Questionnaire for Visual Dysfunctions (SQVD) was developed. We started with a pilot version of 33 items on which a psychometric reduction of items was made to obtain the final questionnaire.

The study followed the tenets of the Declaration of Helsinki, and informed consent was obtained from all subjects after explanation of the nature of the study. If patient was under 18 years old, informed consent was obtained from the parent and/or legal guardian. This research was approved by the ethical committee of the University of Alicante.

The development of the SQVD was based on a 33-item pilot questionnaire which was developed according to scientific evidence. García-Muñoz et al.³ reported the symptoms described by patients in the scientific literature (by means of clinical history or questionnaires) when considering any type of visual anomalies and grouped them into 34 categories of symptoms. Using a Delphi methodology, it was discussed the inclusion of these 34 different categories of visual symptoms in a questionnaire of visual symptomatology, for which experts considered to add other symptom categories¹⁸ so that these results were the origin of this pilot questionnaire.

Thus, considering the Delphi results, and to ensure the validity for the content development aspect and seek patient input in that content development phase, we created a cognitive pre-test²² which had 47 questions of multiple choice answers on a Likert scale of 4 answers (No, Occasionally, Often and Almost always), which measured the frequency of each item. This preliminary instrument with 47 questions was evaluated in a clinical sample of 30 patients with ages between 18 and 71 years old (mean age 29.83 ± 11.07 years old) from an optometric clinic.

In addition to the 47 items of the instrument, a semi-structured interview was conducted, where the patients were asked about those 47 questions that the initial questionnaire had. In this way, they were asked if the questionnaire was too long, if the way to answer the questions was well understood, if the mode of representing the frequency of the response was well understood, or if there were any questions that they had not understood (specifying which one). In addition, they were asked about those questions that we initially thought they could say the same, to see what patients felt. We also asked if they believed it would be better understood with other words to ask about frequency. And also, at the end, patients were asked if they wanted to give their opinion or any suggestion about the items.

Once obtained all of this information, a first qualitative reduction of items was done, eliminating repeated questions and making the changes suggested by the patients, so that we obtained the 33-item pilot scale. These 33 items were randomly distributed into the pilot instrument. They were written as multiple-choice questions on a Likert scale of 4 responses which measured the frequency of the item. The meaning of each answer (which was described in the instructions of the questionnaire) was the following:

- No: the symptom never occurs.
- Occasionally: the symptom occurs at least once every 15 days.
- Often: the symptom occurs once or twice a week.
- Almost always: the symptom occurs almost every day.

This pilot instrument was carried out on a consecutive population sample of 125 patients of an optometric clinic, aged between 15 and 84 (41.72 ± 14.92 years old). Regarding the inclusion criteria, subjects without any type of ocular pathology, dry eye were included. Of the clinical sample, 74 people (59.2%) were women and 51 (40.8%) men. 68 patients (54.4%) were non-presbyopic and 57 (45.6%) were presbyopic. According to visual anomalies, 95 patients had refractive dysfunctions, 18 binocular dysfunctions, 7 accommodative anomalies and 5 subjects did not have any visual anomaly.

Rasch analysis. Rasch analysis is a probabilistic model that assumes items vary in difficulty²³. It estimates the difficulty of items (*item difficulty*) and the relative abilities of the persons (*person ability*) associating them in

a common invariant interval-level scale so that this allows an easy comparison of measures²⁴. It transforms the ordered qualitative observations into quantitative interval level measures on a linear logit scale²⁵. Item difficulty refers to the difficulty level of each item relative to other items in the scale²⁶, so negative values indicate less severity and positive values show greater severity. Person ability refers to the ability level of each person relative to other persons in the sample²⁶.

Rasch analysis was carried out using Winsteps software (version 4.8.1). Rasch parameters analyzed included: functionality of the response categories, fit statistics, local dependency, differential item functioning (DIF), person and item reliability, unidimensionality, targeting and transformation table²⁷.

As items were polytomous, a choice was needed between the Partial Credit Model (PCM) or the Andrich Rating Scale Model (ARSM) for Rasch analysis. PCM considers a different rating scale for each item, while ARSM assumes equal category thresholds across items. We applied the ARSM for polytomous responses after testing it against the PCM using the likelihood ratio statistic. As the test was not significant ($p > 0.05$), the ARSM was appropriate as all items shared the same rating scale²⁸.

Functionality of the response categories. Rasch analysis also provides information about the best number of response categories in the scale. In order to analyze whether the category calibration increases in an orderly manner, response options were assessed with category probability curves (CPC)²⁹. They show the likelihood that a subject with a specific person measure relative to item difficulty will select the category²⁰. The threshold is the midpoint between adjacent response categories so that it reveals the point where the likelihood of choosing either response category is the same³⁰. If a disordered threshold occurs, the situation must be amended collapsing the category needed into an adjacent category^{20,31,32}. For detecting this situation, the Andrich threshold measure must be examined so that thresholds should be spaced at least 1.4 logits²⁹.

Fit statistics. Rasch fit statistics, infit and outfit mean square are obtained to explore the compatibility of the data with the model^{23,33}. They compare the predicted responses to those observed. *Infit* refers to information-weighted fit. It is more sensitive to the pattern inlying observations and less sensitive to outliers. *Outfit* indicates outlier-sensitive fit. It is more sensitive to outliers (atypical cases) so that allows to detect unusual events that occur unexpectedly. Infit and outfit mean squares (MNSQ) closer to 1 indicate a good fit to the model. They must have values between 0.70 and 1.30 logit range³⁰. Values less than 0.70 indicate possible redundancy of items and values greater than 1.30 suggest that items may be measuring something different to the overall scale.

Accordingly, infit and outfit mean square were used to develop the item reduction of the initial instrument and when the reduction was done, the overall infit and outfit mean square of the final scale were obtained to test if the data fit the Rasch model for knowing if the model fit adequate, poor or excellent.

Item reduction criteria were applied using the guidelines described by several authors^{20,23,30,34}. According to these criteria, items were eliminated using the following order of priority:

1. Items with values infit mean square outside 0.70 to 1.30.
2. Items with outfit mean square outside 0.70 and 1.30.
3. Items with a high proportion of missing data (> 50%).
4. Items with ceiling effect: a high proportion of responses in item end-response category (> 50%).
5. Items with a considerably different standard deviation of scores to other items.
6. Items with coefficients of skewness and kurtosis outside the range + 2 to - 2.

Item reduction was done by means an iterative process so that one item was removed at a time²³. Thus, when an item was removed, fit to the model was reestimated consequently as it has been shown that fit is relative so that removal of items leads to variations in fit. Then, the item with the highest number of candidate criteria, ordered by priority was removed first.

Local dependency. Local dependency determines whether the response to any item has a direct impact on the response to any other item³⁵. Local independence is a requirement of the model so that the dependence of items is analyzed to detect its violation with the Yen's Q3 statistic. Hence, local dependency was examined using the residual correlation matrix. As it is stated by Christensen et al.³⁵, no single critical value can be considered to indicate dependency. Simulations have shown that the Q3 critical value appears to be reasonably stable around a value of 0.2 above the average correlation. Therefore, it has been shown that any residual correlation > 0.2 above the average correlation would appear to indicate local dependency.

Accordingly, once performed the reduction of items, it was analyzed the local dependency to determine if there were redundant items which had also to be eliminated. If a pair of items had local dependency (when residual correlations exceeding the mean of all residual correlation by 0.20), items were removed or retained so that those items with an outfit MNSQ closed to 1 must be retained³⁶.

Differential item functioning. Differential item functioning (DIF) is an analysis used to determine if an item measures a latent construct in the same way for different groups³⁴. It represents the differences in item difficulty between respondent groups and it is important to assess it as DIF can affect fit to the model and may damage measures³⁷. In this study DIF analysis was evaluated for gender and presbyopia (presbyopes and non-presbyopes subjects, considering presbyopia when the patient needed addition). Several authors have stated that mean differences in person measures between compared groups should be less than 1.0 logit²³ so that when the value is greater than 1.0 logits, a notable DIF must be considered³⁷. Accordingly, a statistically significant ($p < 0.05$) DIF

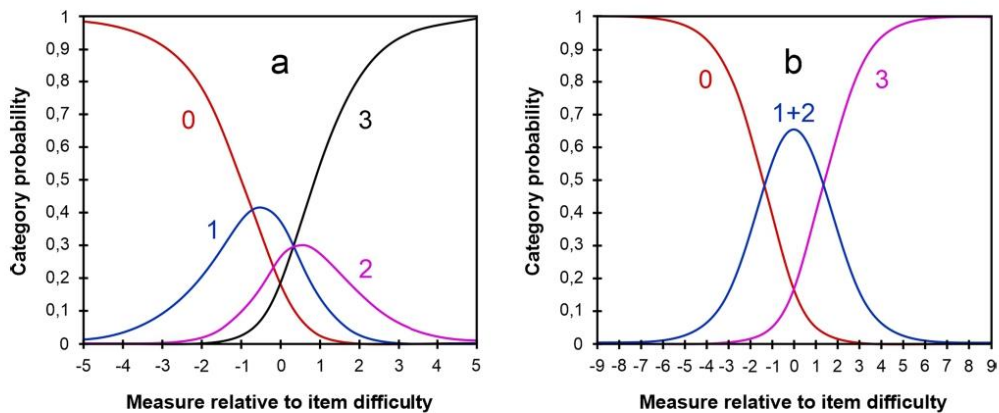


Figure 1. (a) Shows the category probability curves (CPC) for the instrument with 33 items and four response categories. Each curve in the CPC graph represents one response category (No = 0; Occasionally = 1; Often = 2; Almost always = 3). The point where two adjacent curves overlap is the threshold. At this intersection, it is the same likelihood of choosing one category or another. (b) CPC for the instrument with 33 items collapsing the categories 1 (Occasionally) and 2 (Often).

contrast with a difference of more than 1.0 logit was considered for having DIF²⁴. The Rasch-Welch t test method was used to determine the significance of the DIF contrast.

Person and item reliability. Rasch analysis also provides reliability (separation index) for both, person and item, showing the overall performance of the instrument²⁰.

Reliability (separation index) means reproducibility of relative measure location, so that person and item reliability determine the replicability of the person and item locations along the trait continuum²⁴. Accordingly, high reliability (of persons or items) means that there is a high probability that persons (or items) rated with high measures actually do have higher measures than persons (or items) estimated with low measure.

Person separation index is used to classify people, so that a low person separation implies that the instrument may not be enough sensitive to distinguish between high and low performs, showing that more items are needed. Similarly, item separation index is used to verify the item hierarchy, so that a low item separation implies that the sample is not large enough to prove the item difficulty hierarchy of the scale.

Thus, it has been suggested that person and item reliability can range between 0 and 1, which high values indicating better reliability. For person reliability a value of more than 0.80 is acceptable (person separation index > 2 logits) which it means that the measure can stratify the population into at least three groups based on their latent trait measure³⁸. For item reliability a value of more than 0.90 (item separation index > 3 logits) is considered acceptable^{20,39}.

Unidimensionality. Unidimensionality refers to the assumption that the items summed all together form a unidimensional scale²⁷. It was explored by means of the principal component analysis (PCA) of the residuals, considering residuals the differences between observed data and the estimation of the model. The magnitude of the first contrast of the residual is an important indicator, so that this result should not be above 2 eigenvalue²⁷.

Targeting. Rasch analysis also examines the targeting by means of the person-item map⁴⁰. Targeting is the difference between the person ability mean and the item difficulty mean. The closer the person ability mean is to the item difficulty mean, the better the targeting. A difference of zero between both values will indicate a perfect targeting of the scale and a difference of more than 1 logit indicates mistargeting⁴⁰.

Transformation table. Once a best fitting model has been found, Rasch analysis also allows to transform the ordinal scores of the questionnaire to an interval scale. This conversion table was obtained considering ordinal scores range of the scale. Then, the corresponding interval-level scores in logits and ordinal scale range were obtained.

Results

Figure 1a shows the CPC for the 33-item pilot SQVD with its four initial response categories. The threshold shown by the response categories 1 and 2 and the threshold of categories 2 and 3 corresponds to the same value of the person measure relative to item difficulty. This implies that the category calibration does not increase adequately, the thresholds are disordered, so that this situation had to be amended. For that, the Andrich thresholds were examined (Table 1). It shows that for the four initial categories, these values were not within the value

	Category	Andrich thresholds
A	0	None
	1	-0.67
	2	0.25
	3	0.42
B	0	None
	1+2	-1.34
	3	1.34
C	0	None
	1	-0.37
	2+3	0.37

Table 1. Andrich thresholds values (logits) for the 33-item pilot version of SQVD with its original four categories and collapsing categories. A: results for 33-item pilot version with the initial four categories of response (0–1–2–3). B: results when collapsing the categories 1 and 2. C: results when collapsing categories 2 and 3.

Residual correlation	Item	OUTFIT MNSQ	Result	Item	OUTFIT MNSQ	Result
0.41	31	0.88	Removed	33	1.01	Retained
0.37	28	0.92	Removed	33	1.01	Retained
0.34	25	0.95	Retained	31	0.88	Removed
0.33	28	0.92	Removed	31	0.88	Removed

Table 2. Local dependency analysis by means of Yen's Q3 statistic. Pair of items with a residual correlation > 0.2 above the average correlation (0.12) indicate dependency. In this case, residual correlation greater than 0.32 between them were shown. Items with an outfit MNSQ closed to 1 were retained. (MNSQ: mean square).

of 1.4 logits between thresholds indicating that this premise is verified when categories 1 and 2 were collapsed. Accordingly, initial categories 1 (Occasionally) and 2 (Often) were collapsed into the category *Occasionally / Often*, resulting three category response options: 0 (No), 1 (Occasionally / Often) and 2 (Almost always). Figure 1b shows the CPC for the 33-item pilot version of SQVD with these three response categories.

Once the instrument had collapsed into three categories, the item reduction was analyzed. Using the criteria explained in the section of methods for reducing the instrument, 17 items were removed, and were the following: 3, 4, 5, 7, 9, 10, 11, 13, 16, 18, 20, 21, 22, 24, 27, 29, 30. Once removed these items, a version of 16 items was obtained.

With this version of the scale, the local dependency was performed to test if more items had to be removed. As the average correlation was 0.12, those residual correlation values > 0.32 (0.12 + 0.2) indicated dependency. Table 2 shows the residual correlation values greater than the value of 0.32 for the pair of items indicating dependency. It also indicates those items which were removed and retained using the local dependency criteria to reduce items. Once applying these criteria, items 28 and 31 had to be removed, so that the final questionnaire was comprised with 14 items. With these 14 items the local dependency was again analyzed, and no local dependency was obtained.

Table 3 shows the infit and outfit mean squares for the 14 items of the final instrument. It also shows DIF contrast results for the variables examined, for which there were no significant statistically differences ($p > 0.05$) regarding gender for all items (< 1 logits). With respect to the analysis of presbyopia, there were no significant statistically differences for all items (< 1 logits), except for the item 33.

The overall infit and outfit mean square of the final scale are shown in Table 4, in which it can also be seen the person and item reliability values.

When exploring principal component analysis (PCA) of the residuals, the magnitude of the first contrast of the residual was 1.908 eigenvalue, showing the unidimensionality of the instrument.

According to the targeting, the person-item map is shown in Fig. 2. Left-hand column indicates person ability (in logits), where those patients with higher ability are shown at the top of the figure. The mean person ability (shown in the figure as the left M) has a value of -1.59 logits. The right-hand column indicates the item difficulty, for which the mean (right M) is always 0. Accordingly, more difficult items are on the top of the figure. With these results, the targeting of the scale was -1.59 logits.

Transformation table results are shown in Table 5. It shows the ordinal scale and the corresponding interval-level scores. The ordinal scores range from 0 to 28 and corresponding interval scores in logits are included. Interval scores in logit units were rescaled into the ordinal scale range (0–28).

Item	Infit MNSQ	Outfit MNSQ	Gender DIF contrast	Presbyopia DIF contrast
1	0.99	0.96	0.37	0.33
2	1.04	0.97	0.32	0.13
6	0.98	0.91	0.29	0.49
8	0.98	0.95	0.00	0.08
12	1.01	1.00	0.53	0.10
14	0.96	1.06	0.33	0.89
15	1.03	1.06	0.37	0.71
17	0.78	0.74	0.37	0.00
19	1.19	1.24	0.19	0.53
23	0.86	0.85	0.19	0.05
25	1.10	0.99	0.36	0.21
26	0.98	1.07	0.24	0.13
32	0.84	0.95	0.13	0.00
33	1.18	1.13	0.07	1.71*

Table 3. Item Rasch analysis results of the SQVD. (MNSQ: Mean square statistics; DIF: Differential item functioning). * $p < 0.05$.

Persons		Items		Reliability (Separation index)	
Infit MNSQ	Outfit MNSQ	Infit MNSQ	Outfit MNSQ	Person reliability (Separation)	Item reliability (Separation)
0.99	0.99	0.99	0.99	0.81 (2.11)	0.80 (2.06)

Table 4. Summary of the global fit statistics for person ability and item difficulty parameters for the SQVD. (MNSQ: Mean square statistics).

Discussion

Results of this study show that the SQVD has acceptable psychometric properties in a clinical population. Based on a previous research using Delphi methodology, the 33-item pilot instrument was reduced by Rasch analysis until the final version of 14 items was achieved. This final version of SQVD showed adequate fit statistics to the Rasch model, without local dependency or significant statistically DIF for gender and presbyopia, and an adequate reliability and unidimensionality.

The 33-item pilot version showed an inadequate order in the response categories, so it was necessary to collapse category 1 (occasionally) and 2 (often). Therefore, the SQVD response categories were reduced to three: no / occasionally-often / almost always, which allowed ordered thresholds with separations greater than 1.4 logits between them. This situation of collapsing categories has also been used by several questionnaires related to vision^{16,17,25}.

According to the reduction of the scale, 19 items from the 33-item pilot version of SQVD were removed. These 19 items were related to reading problems, visual fatigue, diplopia, ocular problems, blurred vision and postural problems. Accordingly, the final version SQVD had 14 items, a number of items similar to those questionnaires showed in the scientific literature related to visual symptoms^{5,10–12}. The 14 items that were finally included in the SQVD were related to blurred vision, binocular vision problems, ocular irritation, concentration difficulties, reading problems and headache.

All these 14 individual items demonstrated a satisfactory fit to the Rasch model, which could be considered productive for measurement (infit MNSQ between 0.78 and 1.19; outfit MNSQ between 0.74 and 1.24). Mean infit and outfit MNSQ for both subjects (mean infit 0.99 and outfit 0.99 logits) and items (mean infit 0.99 and outfit 0.99 logits) are within the range suggested by scientific literature^{20,23,29,34,41}, so it can be considered that there is an adequate fit to the model. Furthermore, differential item functioning analysis showed that the SQVD had not significant statistically DIF by gender and presbyopia. There was only one item, item 33, which showed significant DIF for presbyopia. As this item is related to blurred near vision, it could argue that this symptom is more specific of a presbyopic population, so we consider that this item should not be removed from the questionnaire. In any case, DIF results allow to assume that the model and the set of item parameters are similar for all comparable groups.

According to the reliability, the person reliability result of 0.81 (person separation 2.11) was satisfactory, which implies the good internal consistency for the SQVD in this sample. The instrument could stratify the population into at least 3 groups based on the latent trait measured. However, the low item reliability of 0.80 (item separation 2.06) shows that it would be necessary a greater sample to prove the item difficulty hierarchy of the scale.

The basic assumption of the Rasch model¹⁷ was demonstrated for the SQVD as the principal component analysis (PCA) results of 1.908 eigenvalue proved the unidimensionality of the scale.

Targeting result of -1.59 logits shows that the SQVD 14 item version has poor targeting, revealing a floor effect, situation clearly exhibited in Fig. 2 where it can be seen the important number of patients who are situated

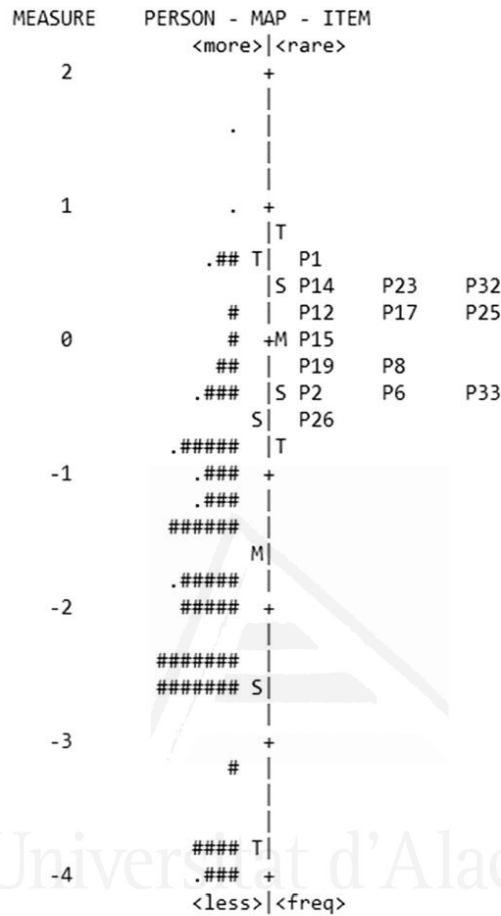


Figure 2. Person-item map for the SQVD. Patients are represented on the left of the dashed line by the symbol “#” (which represents 2 subjects) and “.” (which indicates 1 subject). On the right of dashed line are illustrated the items of SQVD 14-item version with their number (Pnumber of item). M indicates the mean measure (on the left the person ability and on the right the item difficulty). S shows one standard deviation from the mean and T denotes two standard deviations. Higher ability for persons (higher frequency of symptoms) and more difficult items are on the top of the figure.

on the bottom of the figure. This might be explained by overall lower symptoms in the current sample, showing that in this sample, there is a greater number of subjects with less symptomatology. Other authors have shown that this situation is common in questionnaires related to symptoms, as many patients may not have symptoms^{17,21} or due to there is a tendency for subjects to underreport their discomfort¹⁶. These conditions are applicable to the SQVD, so accordingly, the targeting may be considered reasonable. In any case, it would be desirable for future studies to analyze samples that have higher levels of symptoms.

In this study it has been shown the conversion from ordinal to interval-level data²⁷ without modifying the original response of the instrument. Using the conversion table provided here, users are able to increase the precision of the SQVD and thus confidence in any reported finding. It may be useful for clinical purposes, for example when consider reporting changes of the measure variable, as equal interval scaling allows to detect any variations. In addition to that, the conversion to interval scores facilitates the use of parametric statistics. However, it only may be used when there are no missing data, that is, when the completed data for all items are available for the assessed person.

Ordinal measure	Interval measure	
	Logit	Scale
0	-5.19	0.00
1	-3.93	3.39
2	-3.17	5.45
3	-2.69	6.75
4	-2.32	7.74
5	-2.01	8.57
6	-1.74	9.30
7	-1.50	9.96
8	-1.27	10.59
9	-1.04	11.18
10	-0.83	11.76
11	-0.62	12.33
12	-0.41	12.89
13	-0.21	13.45
14	0.00	14.00
15	0.21	14.55
16	0.41	15.11
17	0.62	15.67
18	0.83	16.24
19	1.04	16.82
20	1.27	17.41
21	1.50	18.04
22	1.74	18.71
23	2.01	19.43
24	2.32	20.26
25	2.69	21.25
26	3.17	22.55
27	3.93	24.61
28	5.19	28.00

Table 5. Converting from a raw SQVD score (0–28) to an interval scale in logit units and using the original scale metrics.

The SQVD is difficult to compare with other visual symptoms questionnaires due to the differences in their purposes. Thus, the Conlon Visual Discomfort Scale Survey questionnaire⁵ is an instrument to measure the symptoms of visual discomfort, but this is different to visual symptoms related to the visual anomalies. Visual discomfort may be due to any cause, including for example the glare produced by a light source, so it is not comparable to the SQVD. Similarly, the CISS^{10–12} is a specific questionnaire only for a binocular anomaly (convergence insufficiency), so it cannot be applied to the others visual anomalies and therefore its categories of symptoms cannot be compared to those of the SQVD. Furthermore, the CISS instrument has not been validated using the Rasch model.

This study had some limitations. Although the pilot version of 33 items was distributed to 125 subjects, the sample size should be considered small. Several authors⁴² have suggested that the number of subjects could vary from 4 to 10 subjects per item, with a minimum of 100 subjects. However other authors⁴³ have shown that there are not absolute rules for the sample size needed to validate a questionnaire.

However, the study has several strengths. One of them is related to the population characteristics of the sample. From a clinical point of view, an instrument as the SQVD has practical benefits for clinical purposes as it may be an aid to calibrate the patients' symptoms who attend in an optometric clinic. Because the clinical population should be the target population of this instrument, it is a strength that its development had been done using a clinical sample. Furthermore, the starting point of the study is also another strength to consider. Thus, the meticulous process developed to determine which symptoms should be included in a questionnaire of visual symptoms (the previous Delphi method used)¹⁸, the detailed procedure developed to design the initial scale, including a previous questionnaire in other clinical sample and the comprehensive patient consultations, contributed to the strength of the study.

In conclusion, the SQVD is an instrument with 14 items which shows that data adequately fit the Rasch model, without local dependency or significant statistically DIF for gender and presbyopia, and an adequate reliability and unidimensionality. This allows to this instrument to have practical benefits not only for clinicians but researchers as it could be used to monitor symptoms in subject who are attended in a clinical center with

any type of visual anomalies and supervise their treatment. Future studies should be done to confirm the SQVD behavior in a larger sample with patients more symptomatic.

Received: 13 January 2021; Accepted: 7 July 2021

Published online: 21 July 2021

References

- Cacho-Martínez, P., García-Muñoz, A. & Ruiz-Cantero, M. T. Is there any evidence for the validity of diagnostic criteria used for accommodative and nonstrabismic binocular dysfunctions?. *J. Optom.* **7**, 2–21 (2014).
- Cacho-Martínez, P., Cantó-Cerdán, M., Carbonell-Bonet, S. & García-Muñoz, A. Characterization of visual symptomatology associated with refractive, accommodative, and binocular anomalies. *J. Ophthalmol.* **2015**, 895803 (2015).
- García-Muñoz, A., Carbonell-Bonet, S. & Cacho-Martínez, P. Symptomatology associated with accommodative and binocular vision anomalies. *J. Optom.* **7**, 178–192 (2014).
- Russell, G. E. & Wick, B. A prospective study of treatment of accommodative insufficiency. *Optom. Vis. Sci.* **70**, 131–135 (1993).
- Conlon, E. G., Lovegrove, W. J., Chekaluk, E. & Pattison, P. E. Measuring visual discomfort. *Vis. Cogn.* **6**(6), 637–663 (1999).
- Birnbaum, M. H., Soden, R. & Cohen, A. H. Efficacy of vision therapy for convergence insufficiency in an adult male population. *J. Am. Optom. Assoc.* **70**, 225–232 (1999).
- Adler, P. Efficacy of treatment for convergence insufficiency using vision therapy. *Ophthalmic Physiol. Opt.* **22**, 565–571 (2002).
- Gallaway, M., Scheiman, M. & Malhotra, K. The effectiveness of pencil pushups treatment for convergence insufficiency: A pilot study. *Optom. Vis. Sci.* **79**, 265–267 (2002).
- Borsting, E. et al. Association of symptoms and convergence and accommodative insufficiency in school-age children. *Optometry* **74**, 25–34 (2003).
- Borsting, E. et al. Validity and reliability of the revised convergence insufficiency symptom survey in children aged 9 to 18 years. *Optom. Vis. Sci.* **80**, 832–838 (2003).
- Rouse, M. et al. Validity and reliability of the revised convergence insufficiency symptom survey in adults. *Ophthalmic Physiol. Opt.* **24**, 384–390 (2004).
- Borsting, E., Rouse, M. & De Land, P. Prospective comparison of convergence insufficiency and normal binocular children on CIRS symptom surveys. Convergence insufficiency and reading study (CIRS) group. *Optom. Vis. Sci.* **76**, 221–228 (1999).
- Sterner, B., Gellerstedt, M. & Sjöström, A. Accommodation and the relationship to subjective symptoms with near work for young school children. *Ophthalmic Physiol. Opt.* **26**, 148–155 (2006).
- Vaughn, W., Maples, W. C. & Hoenes, R. The association between vision quality of life and academics as measured by the college of optometrists in vision development quality of life questionnaire. *Optometry* **77**, 116–123 (2006).
- Rouse, M. et al. Academic behaviors in children with convergence insufficiency with and without parent-reported ADHD. *Optom. Vis. Sci.* **86**, 1169–1177 (2009).
- González-Pérez, M., Susi, R., Antona, B., Barrio, A. & González, E. The computer-vision symptom scale (CVSS17): Development and initial validation. *Invest. Ophthalmol. Vis. Sci.* **55**, 4504–4511 (2014).
- Seguí, M. M., Cabrero-García, J., Crespo, A., Verdú, J. & Ronda-Pérez, E. A reliable and valid questionnaire was developed to measure computer vision syndrome at the workplace. *J. Clin. Epidemiol.* **68**, 662–673 (2015).
- Cantó-Cerdán, M., Cacho-Martínez, P. & García-Muñoz, A. Delphi methodology for symptomatology associated with visual dysfunctions. *Sci. Rep.* **10**, 19403 (2020).
- Cappelleri, J. C., Lundy, J. J. & Hays, R. D. Overview of classical test theory and item response theory for the quantitative assessment of items in developing patient-reported outcomes measures. *Clin. Ther.* **36**, 648–662 (2014).
- Pesudovs, K., Burr, J. M., Harley, C. & Elliott, D. B. The development, assessment, and selection of questionnaires. *Optom. Vis. Sci.* **84**, 663–674 (2007).
- McAlinden, C., Pesudovs, K. & Moore, J. E. The development of an instrument to measure quality of vision: The quality of vision (QoV) questionnaire. *Invest. Ophthalmol. Vis. Sci.* **51**, 5537–5545 (2010).
- Perneger, T. V., Courvoisier, D. S., Hudelson, P. M. & Gayet-Ageron, A. Sample size for pre-tests of questionnaires. *Qual. Life Res.* **24**, 147–151 (2015).
- Pesudovs, K., Garamendi, E., Keeves, J. P. & Elliott, D. B. The activities of daily vision scale for cataract surgery outcomes: Re-evaluating validity with Rasch analysis. *Invest. Ophthalmol. Vis. Sci.* **44**, 2892–2899 (2003).
- Kandel, H. et al. Psychometric properties of the keratoconus outcomes research questionnaire: A save sight keratoconus registry study. *Cornea* **39**, 303–310 (2020).
- McAlinden, C. & Lipson, M. Orthokeratology and contact lens quality of life questionnaire (OCL-QoL). *Eye Contact Lens* **44**, 279–285 (2018).
- Stout, J. L. et al. Rasch analysis of items from two self-report measures of motor function: Determination of item difficulty and relationships with children's ability levels. *Dev. Med. Child. Neurol.* **54**, 443–450 (2012).
- Leung, Y. Y., Png, M. E., Conaghan, P. & Tennant, A. A systematic literature review on the application of Rasch analysis in musculoskeletal disease—A special interest group report of OMERACT 11. *J. Rheumatol.* **41**, 159–164 (2014).
- Masters, G. N. A Rasch model for partial credit scoring. *Psychometrika* **47**, 149–174 (1982).
- Linacre, J. M. Optimizing rating scale category effectiveness. *J. Appl. Meas.* **3**, 85–106 (2002).
- McAlinden, C., Khadka, J., de Paranhos, J. F. S., Schor, P. & Pesudovs, K. Psychometric properties of the NEI-RQL-42 questionnaire in keratoconus. *Invest. Ophthalmol. Vis. Sci.* **53**, 7370–7374 (2012).
- Andrich, D. An expanded derivation of the threshold structure of the polytomous rasch model that dispels any “threshold disorder controversy”. *Educ. Psychol. Meas.* **73**, 78 (2013).
- Andrich, D. The legacies of R. A. Fisher and K. Pearson in the application of the polytomous rasch model for assessing the empirical ordering of categories. *Educ. Psychol. Meas.* **73**, 553 (2013).
- Linacre, J. What do infit and outfit, mean-square and standardized mean?. *Rasch Meas. Trans.* **16**(2), 878 (2002).
- Pesudovs, K., Garamendi, E. & Elliott, D. B. The quality of life impact of refractive correction (QIRC) questionnaire: Development and validation. *Optom. Vis. Sci.* **81**, 769–777 (2004).
- Christensen, K. B., Makransky, G. & Horton, M. Critical values for Yen's Q3: Identification of local dependence in the Rasch model using residual correlations. *Appl. Psychol. Meas.* **41**, 178–194 (2017).
- Yasin, R. M., Yunus, F. A. N., Rus, R. C., Ahmad, A. & Rahim, R. B. Validity and reliability learning transfer item using Rasch measurement model. *Proc.—Soc. Behav. Sci.* **204**, 212–217 (2014).
- Lundstrom, M. & Pesudovs, K. Catquest-9SF patient outcomes questionnaire: Nine-item short-form Rasch-scaled revision of the Catquest questionnaire. *J. Cataract. Refract. Surg.* **35**, 504–513 (2009).
- Kandel, H., Khadka, J., Watson, S. L., Fenwick, E. K. & Pesudovs, K. Item banks for measurement of refractive error-specific quality of life. *Ophthalmic Physiol. Opt.* **41**, 591–602 (2021).
- Bond, T. *Applying the Rasch Model: Fundamental Measurement in the Human Sciences* (Routledge, 2015).

40. Stelmack, J. *et al.* Use of Rasch person-item map in exploratory data analysis: A clinical perspective. *J. Rehabil. Res. Dev.* **41**, 233–241 (2004).
41. Wright, B. D. & Linacre, J. M. Observations are always ordinal; measurements, however, must be interval. *Arch. Phys. Med. Rehabil.* **70**, 857–860 (1989).
42. Terwee, C. B. *et al.* Quality criteria were proposed for measurement properties of health status questionnaires. *J. Clin. Epidemiol.* **60**, 34–42 (2007).
43. Tsang, S., Royse, C. F. & Terkawi, A. S. Guidelines for developing, translating, and validating a questionnaire in perioperative and pain medicine. *Saudi J. Anaesth.* **11**, S80 (2017).

Author contributions

M.C.C.: conceived and designed the study, analysed and interpreted the data. P.C.M.: conceived and designed the study, analysed and interpreted the data, and drafted the manuscript, which was revised by all the authors. F.L.L.: conceived and designed the study and carried out the questionnaire. A.G.M.: conceived and designed the study, analysed and interpreted the data. All the authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to P.C.-M.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021

Universitat d'Alacant
Universidad de Alicante

2.3. CAPÍTULO III: VALIDATION OF THE SYMPTOM QUESTIONNAIRE FOR VISUAL DYSFUNCTIONS (SQVD): A QUESTIONNAIRE TO EVALUATE SYMPTOMS OF ANY TYPE OF VISUAL DYSFUNCTIONS

Referencia: Cacho-Martínez P, Cantó-Cerdán M, Lara-Lacárce F, García-Muñoz A. Validation of the Symptom Questionnaire for Visual Dysfunctions (SQVD): A Questionnaire to Evaluate Symptoms of any Type of Visual Dysfunctions. *Transl Vis Sci Technol.* 2022;11:7.

Autores:

1. Pilar Cacho-Martínez, PhD.

Departamento de Óptica, Farmacología y Anatomía. Universidad de Alicante.

2. Mario Cantó-Cerdán, MSc.

Departamento de Óptica, Farmacología y Anatomía. Universidad de Alicante.

3. Francisco Lara-Lacárce F, PhD.

Departamento de Oftalmología, Optometría, Otorrinolaringología, y Anatomía Patológica. Universidad de Murcia.

4. Ángel García-Muñoz, PhD.

Departamento de Óptica, Farmacología y Anatomía. Universidad de Alicante.

Validation of the Symptom Questionnaire for Visual Dysfunctions (SQVD): A Questionnaire to Evaluate Symptoms of any Type of Visual Dysfunctions

Pilar Cacho-Martínez¹, Mario Cantó-Cerdán¹, Francisco Lara-Lacárcel², and Ángel García-Muñoz¹

¹ Departamento de Óptica, Farmacología y Anatomía, Universidad de Alicante, Alicante, Spain

² Departamento de Oftalmología, Optometría, Otorrinolaringología y Anatomía Patológica, Universidad de Murcia, Murcia, Spain

Correspondence: Pilar Cacho-Martínez, Departamento de Óptica, Farmacología y Anatomía, Apartado 99, Universidad de Alicante, 03080 Alicante, Spain. e-mail: cacho@ua.es

Received: July 26, 2021

Accepted: January 5, 2022

Published: February 3, 2022

Keywords: visual symptomatology; questionnaires; visual dysfunctions; Rasch analysis; refractive, accommodative, and binocular dysfunctions

Citation: Cacho-Martínez P, Cantó-Cerdán M, Lara-Lacárcel F, García-Muñoz Á. Validation of the symptom questionnaire for visual dysfunctions (SQVD): A questionnaire to evaluate symptoms of any type of visual dysfunctions. *Transl Vis Sci Technol.* 2022;11(2):7. <https://doi.org/10.1167/tvst.11.2.7>

Purpose: To assess psychometric properties of the Symptom Questionnaire for Visual Dysfunctions (SQVD) questionnaire, including accuracy, validity, and reliability, in a clinical sample of patients having any type of visual dysfunction.

Methods: A clinical sample of 306 patients self-administered the SQVD. Rasch analysis was performed to analyze the functionality of the response categories, fit statistics, differential item functioning (DIF), person and item reliability, targeting, local dependency, unidimensionality, and transformation table. Accuracy was assessed by means of receiver operating characteristic (ROC) curves, using symptoms reported in each patient's clinical record as the gold standard for classifying patients with and without symptoms. The concurrent validity, known group validity, and test-retest reliability (repeatability, using the intraclass correlation coefficient [ICC]) were also examined.

Results: SQVD showed orderly category responses. The 14 items fit the Rasch model without significant DIF for gender, presbyopia, and dysfunctions. Person and item reliabilities were 0.81 and 0.85, respectively. Targeting was -1.49 logits. Yen's Q3 statistic showed no local dependency. SQVD was unidimensional (first contrast of the residual = 1.852 eigenvalue with a variance explained by measures of 52.23%). The area under the ROC curve was 0.836 (95% confidence interval [CI], 0.792–0.879) with a cutoff of ≥ 6 showing good accuracy (sensitivity = 0.759; specificity = 0.783). SQVD showed good concurrent and known group validity and high repeatability (ICC, 0.857; 95% CI, 0.710–0.933) when administered twice 1 week apart.

Conclusions: SQVD has shown good psychometric properties. It can be considered an accurate, valid, and reliable questionnaire to detect visual symptoms related to any type of refractive, accommodative, and binocular dysfunction.

Translational Relevance: SQVD may be used for diagnostic purposes, as it can accurately detect symptoms related to any sort of visual dysfunction. It may also be useful to monitor the treatment outcomes of these conditions.

Introduction

Visual dysfunctions (refractive, accommodative, and binocular) are common in clinical practice among patients undergoing visual examinations. These anomalies may cause symptoms that can impact the patient's comfort level when performing visual tasks.¹ The scientific literature has shown that these visual

dysfunctions are related to various symptoms that may be shared by different visual anomalies.^{2,3} However, it has also been shown that there are differences in the way information is collected and in the methodology that is followed to classify anomaly-related symptoms.^{2,3} The most common way to record the symptoms reported by a patient undergoing a visual examination is the patient's clinical record, although several questionnaires related to visual symptoms have



been developed.⁴⁻¹⁵ However, scientific evidence has shown that only three of them have been psychometrically validated.⁶⁻⁹ The Conlon survey was developed to detect visual discomfort,⁷ but it is not related to any specific visual dysfunction,⁷ whereas the Convergence Insufficiency Symptom Survey (CISS) and its version for parents are both specific for convergence insufficiency.^{6,8,9}

Additional questionnaires have also been developed focusing on other visual conditions. Thus, in the literature, we can find scales for patients with low vision,¹⁶ cataracts,^{17,18} keratoconus,¹⁹ or vertical deviations,²⁰ for computer users^{21,22}; and for contact lens wearers,²³ as well as instruments assessing quality of vision²⁴ and quality of life related to various aspects of vision.²⁵⁻²⁸ However, to date, there have been no questionnaires available to capture visual symptoms related to any type of visual dysfunction (refractive, accommodative, or binocular). For this reason, we developed a questionnaire on visual symptomatology referred to as the Symptom Questionnaire for Visual Dysfunctions (SQVD). This process included an initial step in which we compiled from the scientific evidence³ a list of symptoms that could be triggered by these anomalies. Following that systematic review,³ which resulted in a comprehensive list of symptoms related to any type of visual anomalies, we considered including those symptoms in a questionnaire using a Delphi methodology.²⁹ The results of this Delphi study led to the design of the initial scale, which was tested in a small patient sample (including comprehensive patient consultations³⁰) and was used to develop the pilot version of SQVD, which consisted of 33 items. It was tested in a clinical sample of 125 patients and assessed by means of Rasch analysis.³¹ The outcomes of the Rasch analysis showed a reduction in the number of items, which resulted in the final version of SQVD having 14 items to detect the presence and frequency of visual symptoms related to visual anomalies.

As it was necessary to analyze and confirm the appropriateness of this questionnaire in a larger sample, the purpose of the present study was to validate the SQVD in a clinical sample of patients with any type of visual anomalies. The aim was to assess psychometric properties of the SQVD, including accuracy, validity, and test-retest reliability (repeatability).

Methods

Subjects and Procedure

Patients 14 years of age and older consecutively attending visual examinations in a private optomet-

ric clinic were selected as potential participants in this study; they were later recruited if they met all of the inclusion and exclusion criteria. Each patient underwent a complete eye examination. Ocular health and refractive status were evaluated, and accommodative and binocular tests were performed in order to diagnose any type of visual anomaly (be it refractive, accommodative, or binocular).³² The presence of symptoms was confirmed by the optometrist assessment carried out during the case history. The diagnosis of each visual dysfunction relied on the criteria of García-Muñoz et al.³² When the visual examination was completed, those subjects with any type of ocular disease, a history of refractive surgery, or dry eye or who were taking medication that could alter visual function were excluded from the study.

As a result of the abovementioned process, 306 patients between the ages of 14 and 87 years (mean age, 38.38 ± 17.29 years) were included in the study. All of the participants were Spanish, and 184 of them were women (60.1%). Among the participants, 204 patients had some type of visual dysfunction (66.7%): 132 cases of refractive, 14 accommodative (one accommodative insufficiency and 13 accommodative excess), 50 binocular (17 convergence insufficiency, 10 convergence excess, four insufficiency divergence, five basic esophoria, four basic exophoria, one hyperphoria, three amblyopia, and six strabismus), and eight with both accommodative and binocular anomalies (four accommodative excess plus convergence insufficiency; one accommodative excess plus convergence excess; one accommodative excess plus divergence insufficiency; two accommodative insufficiency plus convergence insufficiency). Consequently, 102 subjects in the sample did not show any anomaly (33.3%).

The study was approved by the University of Alicante's Ethics Committee and followed the tenets of the Declaration of Helsinki. Informed consent was obtained from all subjects after giving them an explanation of the nature of the study. For underaged participants (i.e., under 18 years of age), it was their parents or legal guardians who accepted the study's participation principles and signed the informed consent; the participants also gave their assent to participate.

The SQVD was completed by all 306 subjects included in the sample. It was self-administered. Before they undertook their visual examination, patients were given a printed copy of the questionnaire (containing the list of questions and answer options) and were accompanied by a researcher in case they had any questions while answering the questionnaire. This researcher was a different person than the examiner who performed the visual examination, to avoid bias in the study.

The SQVD is shown as Supplementary Material. The tool is presented in its original validated language (Spanish). We have also included its translation into English for non-Spanish-speaking readers, but it must be kept in mind that this is not a cross-cultural adaptation for the English language: the tool in English would also require further cross-validation. The questionnaire has 14 items addressing blurred vision, binocular problems, ocular irritation, headache, concentration difficulties, reading problems, and postural aspects. Each item has three response options (Likert scale) to indicate the frequency of the symptom (propensity of its presence):

- No: the symptom never occurs (0 point).
- Occasionally or often: the symptom occurs sporadically (at least once every 15 days) or once or twice a week (1 point).
- Almost always: the symptom occurs almost every day (2 points).

This explanation is also included in the printed questionnaire. As shown above, the answer for each item is assigned a score between 0 and 2 points, and the total SQVD score is then obtained by adding the 14 individual item scores; thus, scores can range from 0 to 28.

Data Analysis

Rasch analysis was performed and included all parameters that the scientific evidence has shown to be essential when developing and validating a questionnaire.³³ Rasch analysis is a probabilistic model that estimates the difficulty of items (item difficulty) and the relative abilities of the respondents (person ability) and aligns these two parameters in an invariant interval-level scale. It transforms simple ordinal categorical data into interval-level data on a linear logit scale.^{25,34,35} The Andrich Rating Scale Model for Rasch analysis was employed³⁶ using Winsteps 4.8.1. The parameters analyzed included response category functioning, fit statistics, differential item functioning, person reliability, item reliability, targeting, local item dependency, unidimensionality, and transformation table. In addition to the Rasch analysis, a receiver operating characteristic (ROC) analysis was carried out to assess the accuracy of the questionnaire and to obtain the cutoff value for the instrument. The concurrent validity and known group validity were also evaluated; moreover, the repeatability was calculated based on the intraclass correlation coefficient (ICC).

Response Category Functioning

Response options were analyzed by means of category probability curves³⁷ to test whether the response categories were ordered. These curves show the likelihood for a subject to choose a given category. Because the threshold is the midpoint between adjacent response categories, it represents the point where the likelihood of choosing either response category is the same.³⁸ To assess whether the categories are adequately ordered, the scientific literature recommends calculating Andrich thresholds, which should be spaced at least 1.4 logits apart.³⁷ If this situation does not occur, thus showing a disordered threshold, it may be necessary to collapse adjacent categories.^{33,39–41}

Fit Statistics

Rasch fit statistics (infit and outfit mean square [MNSQ]) were obtained to explore whether the data fit the Rasch model expectations.^{33,35,42} Infit and outfit MNSQ values closer to 1 indicate a good fit to the model; that is, that more difficult items are less likely to be affirmed successfully (and vice versa). Values below 0.70 suggest that there may be a redundancy of items, whereas values above 1.30 could mean that items are measuring something different from the overall scale. For this reason, fit statistics values must range between 0.70 and 1.30 logits.³⁹

Differential Item Functioning

Differential item functioning (DIF) analysis is used to determine whether different subgroups respond differently to particular items.¹⁹ It is important to undertake this analysis, as the presence of DIF may impact the fit of the data to the model and may corrupt measures.³³ In this study, DIF analysis was assessed for gender, presbyopia (presbyopes vs. non-presbyopes, considering presbyopia as the need for near addition), and dysfunction (having vs. not having a visual dysfunction). It had been established that mean differences in person measures between groups should be less than 1.0 logit.²⁵ For values above 1.0 logit, a notable DIF must be considered.¹⁸ Following these criteria, in this study, DIF was considered to be present when the findings showed a statistically significant ($P < 0.05$) DIF contrast and a difference above 1.0 logit.¹⁹ The Rasch–Welch *t*-test method was then used to establish the significance of the DIF contrast.

Person and Item Reliability

To test the overall performance of the instrument, we assessed the SQVD person and item reliability (separation index).^{33,39} Person and item reliability determines the replicability of the person and item locations along the trait continuum.¹⁹ The values can

range between 0 and 1, with higher values indicating better reliability. A person reliability value > 0.80 is considered to be acceptable (person separation index > 2 logits) and implies that the measure can stratify the population into at least three groups.³⁴ Accordingly, because the person separation index is used to classify people, a low person separation value suggests that the instrument may not be sensitive enough to discriminate between high and low performance, thus indicating that more items are needed. For item reliability, a value above 0.90 (item separation index > 3 logits) is considered appropriate.^{39,43} In this case, the item separation index is used to evaluate the item hierarchy, such that a low item separation implies that the sample is not large enough to prove the item difficulty hierarchy of the scale.

Targeting

Targeting refers to the difference between the person ability mean and the item difficulty mean. It can be assessed with Rasch analysis exploring the person-item map,⁴⁴ which depicts the spread and hierarchy of subjects and items; the closer the person ability mean is to the item difficulty mean, the better the targeting. A difference of zero between both values is considered perfect targeting of the scale, whereas a difference greater than 1 logit is considered mistargeting.³⁸

Local Dependency

Local dependency determines whether the response to any item has a direct influence on the response to any other item,⁴⁵ and for that reason local independence of items is a requirement of the model. Yen's Q3 statistic was used to detect local dependence by means of the residual correlation matrix. Christensen et al.⁴⁵ showed that no singular critical value can be appropriate to indicate dependency; however, simulations have proved that the Q3 critical value appears to be reasonably stable around a value of 0.2 above the average correlation. Consequently, this is the recommended reference to use, and any residual correlation > 0.2 above the average correlation may be considered to indicate local dependency.

Unidimensionality

Unidimensionality refers to the assumption that the questionnaire measures a single construct; that is, that the items summed together form a unidimensional scale.³³ Principal component analysis (PCA) of the residuals was used to explore this property. The residuals were considered to be the differences between observed data and model-derived estimates. It has been suggested in the literature that the first contrast of the

residual, should not be above 2 eigenvalue, and the variance explained by the measures must be $> 50\%$.³³

Transformation Table

Using Rasch analysis, it is possible to transform the ordinal scores of a questionnaire into an interval scale without modifying the original responses of the instrument.³³ In the present study, this conversion table was obtained by considering the raw scores of the ordinal measures, and then the corresponding interval-level scores in logits and ordinal scale were obtained. This transformation table helps other authors use the questionnaire and obtain Rasch scores, although the patient populations should be similar to that included in this study.

ROC Analysis

The accuracy of SQVD was assessed by means of ROC curves, sensitivity (S) and specificity (Sp).⁴⁶ ROC curves depict the true positive rate (S) versus the false positive rate ($1 - Sp$) over a range of cutoff values. The overall accuracy of a test can be quantified by means of the area under the ROC curve; thus, the larger the area, the better the test.⁴⁶ In order to analyze the diagnostic validity of the SQVD, a ROC analysis was performed using the symptoms reported by the subjects in their case history as the gold standard and subsequent patient classification as subjects with symptoms and asymptomatic subjects. The ROC analysis relied on the original raw score of the SQVD for both groups of patients (i.e., symptomatic and asymptomatic) and was used to choose the area under the ROC curve and the coordinates of the curve. The cutoff point was chosen by means of a balance between S and Sp.⁴⁶ This cutoff is necessary to take into account when a patient passes or fails the SQVD.

Validity

The concurrent validity and known group validity were evaluated. The concurrent validity represents the correlation level between the questionnaire score and the score of clinical measures. In the present study, the SQVD score was correlated with visual acuity based on the Spearman's rho coefficient value. A correlation between 0.3 and 0.9 is considered to be adequate.⁴⁷ As for the known group validity, which is the extent to which the instrument can discriminate between clinically different groups, it was assessed by analyzing the statistically significant differences found between the two groups of patients: those with visual anomalies and those that did not show any visual dysfunction. The Mann-Whitney *U* test was used with a significance level of 0.05.

Test–Retest Reliability

The test–retest reliability refers to how repeatable the results are when the instrument is administered by the same observer.³⁹ It was quantified by means of the ICC⁴⁸ and corresponding 95% confidence intervals. The literature recommends an ICC value ≥ 0.80 to demonstrate temporal stability when administered in two different periods.^{39,47,48} Considering an expected ICC of 0.85, two measurements, a lowest acceptable ICC of 0.7, a significance level of 0.05, and a power of 80%, the required sample size turned out to be 42 patients.⁴⁹ Thus, in order to assess the SQVD repeatability, 50 randomly chosen participants were administered the questionnaire a second time, 1 week after the first.⁵⁰ ROC curves, validity, and test–retest reliability were performed using SPSS Statistics 20.0 for Windows (IBM, Armonk, NY).

Results

The category probability curves depicted in Figure 1 show good category discrimination and orderly category responses. The Andrich threshold revealed separations greater than 1.4 logits between adjacent categories (see Table 1), which indicates ordered thresholds and implies that each category response had equal probability to be endorsed by the patients.

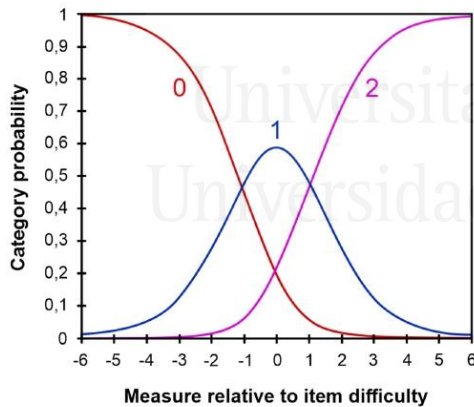


Figure 1. Category probability curves (CPCs) for the SQVD. Each curve in the CPC graph represents one response category (no = 0; occasionally/often = 1; almost always = 2). The figure shows the performance of the three response categories of the SQVD, which asked about the frequency of each of the symptoms under evaluation. The point where two adjacent curves overlap is the threshold, and the two corresponding categories have equal likelihood of being chosen.

Table 1. SQVD Categories and Andrich Threshold Values (Logits)

SQVD Category	Andrich Threshold
0	None
1	-1.34
2	1.34

Table 2 summarizes the infit and outfit MNSQ values for each of the 14 items included in the SQVD. Fit statistics revealed that all 14 items fitted the Rasch model, as they fell within the range values suggested. DIF contrast results are also shown in Table 2. Values greater than 1 logit reveal the presence of DIF. As can be seen in the table, there were no statistically significant DIF differences for gender, dysfunction, or presbyopia (<1 logit) except for items 1 and 2 (related to headache and blurred near vision) which showed DIF for presbyopia. Table 3 shows the overall infit and outfit MNSQ statistics for the SQVD, which indicate that mean infit and outfit (for both subjects and items) fit the Rasch model. Reliability values (person and item reliability) are also shown in Table 3; the person reliability value of 0.81 showed a good person separation index.

With regard to targeting, Figure 2 shows the person–item map for SQVD. Person ability (in logits) is illustrated in the left-hand column; patients with higher ability are shown at the top of the figure. For this patient sample, mean person ability was -1.49 logits (shown in the figure as the left M). The right-hand column displays item difficulty, for which the mean is always 0 (shown as the right M). This poor targeting (-1.49 logits) means that an important number of patients are located at the bottom of the figure, indicating a floor effect. The person–item map reveals that items were targeting the more symptomatic patients; therefore, patients having fewer visual symptoms (especially those at the floor) could not be properly differentiated by the SQVD items.

Yen’s Q3 statistic confirmed that none of the 14 items included in the SQVD showed local dependency, as the residual correlation values were always <0.2 above the average correlation. Furthermore, PCA of the residuals confirmed the unidimensionality of the SQVD, as the magnitude of the first contrast of the residual was 1.852 eigenvalue and the variance explained by the measures was 52.23%.

The transformation table results are provided in Table 4, which shows conversion of the ordinal scale (0 to 28) into the corresponding interval-level scores

Table 2. Rasch Analysis Results for Each SQVD Item

Item	Infit MNSQ	Outfit MNSQ	Gender DIF Contrast	Presbyopia DIF Contrast	Visual Dysfunction DIF Contrast
1	0.93	0.94	0.25	1.15 ^a	0.20
2	1.07	1.06	0.02	1.10 ^a	0.26
3	0.93	0.95	0.00	0.09	0.26
4	1.25	1.25	0.10	0.49	0.00
5	0.89	0.98	0.10	0.14	0.35
6	1.00	0.93	0.15	0.48	0.13
7	0.90	0.91	0.00	0.24	0.45
8	0.90	0.87	0.47	0.21	0.36
9	1.14	1.10	0.58	0.92	0.13
10	1.04	1.05	0.35	0.08	0.51
11	1.05	1.14	0.10	0.00	0.10
12	1.15	1.12	0.52	0.55	0.00
13	0.73	0.70	0.24	0.60	0.47
14	0.96	1.03	0.14	0.80	0.31

^aStatistically significant ($P < 0.05$).

Table 3. Summary of Global Fit Statistics for SQVD Persons and Items, and Reliability Parameters

Persons		Items		Reliability (Separation Index)	
Infit MNSQ	Outfit MNSQ	Infit MNSQ	Outfit MNSQ	Person	Item
1.01	1.00	1.00	1.00	0.81 (2.11)	0.85 (2.41)

(logits) and their subsequent rescaling into the ordinal scale range (0–28) of the tool.

Figure 3 depicts the ROC curves, with an area under the curve of 0.836 ($P < 0.001$; 95% CI, 0.792–0.879). A score cutoff of ≥ 6 had the best balance for S and Sp (0.759 and 0.783, respectively), which means that patient scores ≥ 6 indicate the presence of visual symptoms.

As for concurrent validity, there was a significant correlation (correlation coefficient = 0.246; $P < 0.001$) between logMAR visual acuity and the score obtained in the questionnaire. With regard to known group validity, the mean SQVD score for subjects with some type of visual dysfunction was 8.41 ± 4.25 versus 3.82 ± 2.19 for subjects with no visual dysfunctions, indicating statistically significant differences ($P < 0.001$) between the two groups. If we break down the data by dysfunction type, the mean score for subjects with refractive dysfunctions was 8.61 ± 4.36 , whereas for subjects with accommodative and binocular anomalies it was 8.06 ± 4.07 . There were statistically significant differences between each of these two subgroups and the group of patients with no visual anomalies ($P < 0.001$).

Furthermore, as for test–retest reliability, the ICC was 0.857 ($P < 0.001$; 95% CI, 0.710–0.933).

Discussion

The outcomes of this study confirm that the SQVD has good psychometric properties and is an accurate, valid, and repeatable questionnaire to detect symptoms in patients with any type of visual dysfunction.

The tool has shown that data adequately fit the Rasch model, as all fit statistics fell within the appropriate intervals established by the scientific literature.^{26,35,37,39,51} There was no local dependency for items, and the tool demonstrated unidimensionality.³³

Furthermore, the tool performed similarly for all of the sample subgroups, because, in general, there was no DIF for gender, dysfunction, or presbyopia. The fact that item 1 (headache) and item 2 (blurred near vision) showed DIF for presbyopia indicates that these two symptoms are presbyopia related, thus suggesting a particular symptomatology for this population. Consequently, these two items should not be removed from the questionnaire as they provide information about symptoms of an important visual condition.

The scale is also reliable. The good person reliability value of the SQVD implies that the instrument is able to stratify subjects into at least three groups based on the

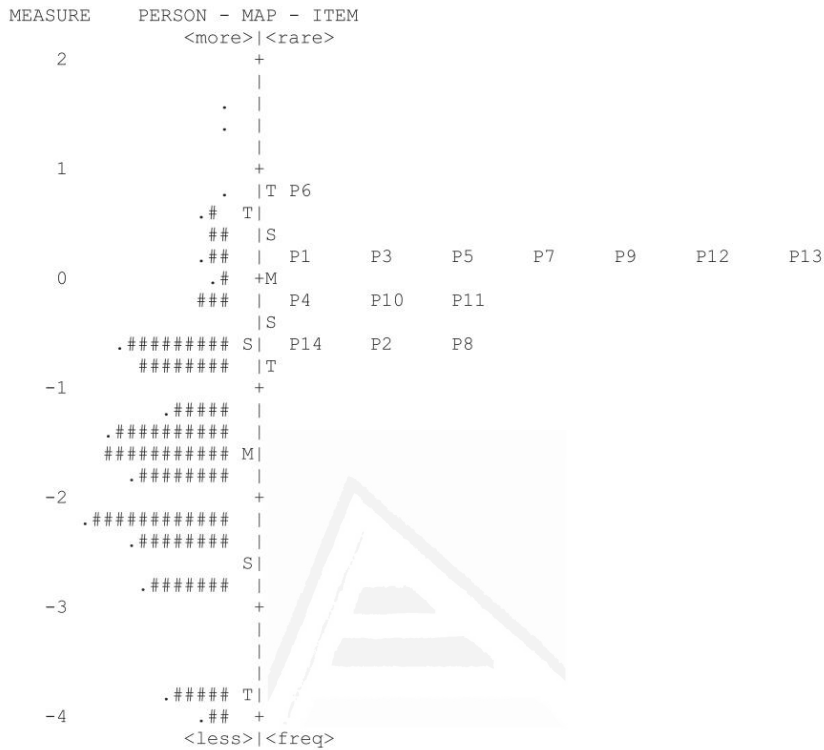


Figure 2. Person-item map for the SQVD. Patients (person ability) are represented on the left of the *dashed line*. The # symbol represents three subjects, and the • symbol indicates one or two subjects. The items of SQVD with their number (Pnumber of item) are shown on the *right of the dashed line* of the figure. “M” indicates the mean measure (*left, person ability; right, item difficulty*). “S” shows 1 SD from the mean, and “T” indicates 2 SDs, all expressed in logits. The person-item map orders the symptoms of the patients and the item difficulty. Higher ability for persons (in this study, higher frequency of symptoms) and more difficult items are at the top of the figure.

translational vision science & technology

latent trait being assessed. However, because the item reliability was 0.85, a larger sample would be required to prove the item difficulty hierarchy of the scale.

The poor targeting by the SQVD, which reveals a floor effect, can be explained by the high number of patients in the sample that reported few symptoms. Similar results have been reported in other studies using different symptom instruments, as many patients say they have no symptoms at all^{22,24} or underreport their discomfort.²¹ Because the same scenario may occur with the SQVD, the targeting may be considered to be reasonable. Future studies could analyze larger samples made up of patients showing higher symptom levels so as to test this targeting.

The SQVD was also shown to be a valid instrument. Concurrent validity outcomes proved that there is a significant statistical correlation between visual acuity

and SQVD score, which suggests that the instrument is able to detect symptoms in these patients. However, the low correlation value implies that these visual anomalies are not always related to a problem affecting visual acuity. For example, although accommodative and binocular anomalies may cause symptoms, they do not necessarily lead to a drop in visual acuity. With regard to known group validity, the instrument has proven to be able to differentiate symptoms of subjects with and without dysfunctions. These outcomes also imply that the SQVD has good validity.

Rasch analysis has also allowed development of a conversion table that transforms ordinal data into interval-level data.³³ With this transformation table, users may improve the precision of the SQVD. For example, a patient with an original ordinal score of 15 will have a corresponding interval score of 14.49,

Validation of the SQVD

Table 4. Conversion of Raw SQVD Scores (0–28) to Interval Scale (Logit Units) Using the Original Scale Metrics

Ordinal Measure, Raw Score	Interval Measure	
	Logit	Scale
0	-4.91	0.00
1	-3.66	3.55
2	-2.91	5.70
3	-2.44	7.04
4	-2.09	8.05
5	-1.80	8.88
6	-1.54	9.60
7	-1.31	10.25
8	-1.10	10.85
9	-0.90	11.42
10	-0.71	11.96
11	-0.53	12.48
12	-0.35	12.99
13	-0.18	13.50
14	0.00	13.99
15	0.17	14.49
16	0.35	15.00
17	0.53	15.51
18	0.71	16.03
19	0.90	16.57
20	1.10	17.14
21	1.31	17.74
22	1.54	18.39
23	1.80	19.11
24	2.09	19.95
25	2.44	20.96
26	2.91	22.30
27	3.67	24.45
28	4.91	28.00

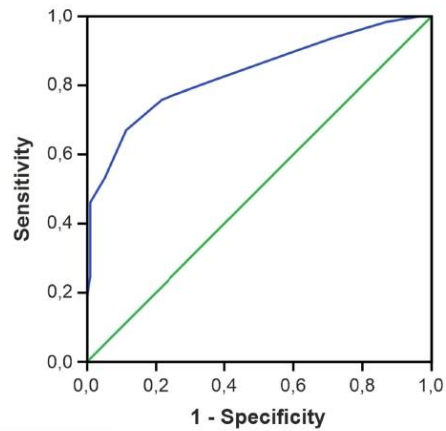


Figure 3. ROC curve for the SQVD questionnaire. The ROC curve summarizes the results of sensitivity and specificity analyses for various cutoff points. The greater the area under the curve, the better the accuracy of the SQVD to detect symptoms.

using the same scale range. By applying this conversion, clinicians can report changes in the variable under evaluation as it better allows use of parametric statistics. The only requirement for applying this conversion is that the patient must complete the entire questionnaire, providing an answer for each and every item.

The SQVD also showed good diagnostic accuracy to detect symptoms related to any type of visual dysfunction. The ROC analysis cutoff value indicates that a score ≥ 6 suggests that the subject's visual symptoms are related to some type of visual dysfunction. Finally, this instrument has also shown high repeatability, which is an important property, for example, when evaluating changes in symptoms when considering a specific treatment.

All of these findings highlight the contributions of this study, as this new scale provides a clinical tool to be used in clinical practice for the exploration of visual anomalies. To the best of our knowledge, no other similar questionnaire can be compared with the SQVD. There is a novel scale for schoolchildren with nonstrabismic binocular vision anomalies,²⁷ but it is designed to quantify quality of life. Even though it has several items that ask about visual symptoms, the goal of the tool is to measure quality of life in a pediatric population, thus comparison with the SQVD would be difficult. The Conlon questionnaire⁷ was developed to quantify visual discomfort, but its description of visual discomfort is neither related nor comparable to symptoms associated with visual dysfunction. Items in the Conlon survey include several symptoms within the same question, which makes it difficult to compare this instrument with other scales. In any case, it does have several questions that are similar to some SQVD items, such as those about headache, blurred vision, ocular irritation, and reading problems. Again, when comparing the SQVD with the CISS survey,⁹ symptoms related to headache, blurred vision, red eyes, the need to reread the text, and sleepy feeling are common to both questionnaires. This could suggest that the CISS could be used to evaluate symptoms due to any type of visual anomaly; however, it was developed for convergence insufficiency only. Although a cross-cultural adaptation for the Spanish language has been made using the Rasch method,⁵² the original questionnaire⁹ was not validated using this method. Several

authors have used the CISS for other visual dysfunctions,^{53–56} but this is not an appropriate approach because a questionnaire should only be used for the specific condition for which it was developed and validated.^{57,58} In fact, other studies in the literature have reported difficulties when applying the CISS to other visual conditions, different from convergence insufficiency.^{59–61}

Such difficulties with existing questionnaires emphasize the need for a survey specifically developed and validated for particular types of visual condition (be it refractive, accommodative, or binocular), such as the SQVD.

The present study does have several limitations. The targeting and item reliability outcomes imply that larger samples of patients with higher symptom levels would be desirable. The pilot study on which the present study was based³¹ analyzed a sample made up 125 subjects; this study, having a larger clinical sample of 306 subjects, has shown improved targeting and item reliability while preserving the reliability and validity of the SQVD. Another study limitation has to do with convergent validity. This property has not been tested, as there is no other questionnaire similar to the SQVD that captures the symptomatology of all visual dysfunctions. However, in future research projects, it would be interesting to test the discriminant validity and compare the SQVD with other instruments that focus on other visual conditions. Despite these limitations, the strength of this study lies in the fact that the recruited patients came from a real-world clinical practice population. For clinical purposes, an instrument that is devised to detect the presence and frequency of symptoms caused by any type of visual dysfunction should be tested in a sample having characteristics similar to those of the population in which it will be applied, which is what we did in the present study. Consequently, this tool will be useful in detecting symptoms in clinical practice, which is the main contribution of this study. The SQVD would be useful not only for diagnostic purposes, as it can accurately detect the presence of symptoms, but also for monitoring symptom severity and frequency in patients who are undergoing treatment for their visual anomalies.

In conclusion, the results of this study show that the SQVD has good psychometric properties and good diagnostic accuracy and is a valid and repeatable questionnaire to detect the presence and frequency of visual symptoms related to any type of visual dysfunction, be it refractive, accommodative, or binocular. Hence, this tool can be used in clinical practice to identify patients with symptoms that are due to visual anomalies and is also valid for research studies.

Acknowledgments

Disclosure: **P. Cacho-Martínez**, None; **M. Cantó-Cerdán**, None; **F. Lara-Lacárcel**, None; **Á. García-Muñoz**, None

References

1. Cacho-Martínez P, García-Muñoz A, Ruiz-Cantero MT. Is there any evidence for the validity of diagnostic criteria used for accommodative and nonstrabismic binocular dysfunctions? *J Optom.* 2014;7:2–21.
2. Cacho-Martínez P, Cantó-Cerdán M, Carbonell-Bonete S, García-Muñoz A. Characterization of visual symptomatology associated with refractive, accommodative, and binocular anomalies. *J Ophthalmol.* 2015;2015:895803.
3. García-Muñoz A, Carbonell-Bonete S, Cacho-Martínez P. Symptomatology associated with accommodative and binocular vision anomalies. *J Optom.* 2014;7:178–192.
4. Russell GE, Wick B. A prospective study of treatment of accommodative insufficiency. *Optom Vis Sci.* 1993;70:131–135.
5. Birnbaum MH, Soden R, Cohen AH. Efficacy of vision therapy for convergence insufficiency in an adult male population. *J Am Optom Assoc.* 1999;70:225–232.
6. Borsting E, Rouse M, De Land P. Prospective comparison of convergence insufficiency and normal binocular children on CIRS symptom surveys. Convergence Insufficiency and Reading Study (CIRS) group. *Optom Vis Sci.* 1999;76:221–228.
7. Conlon EG, Lovegrove WJ, Chekaluk E, Pattison PE. Measuring visual discomfort. *Vis Cogn.* 1999;6:637–663.
8. Borsting E, Rouse M, Mitchell G, et al. Validity and reliability of the revised Convergence Insufficiency Symptom Survey in children aged 9 to 18 years. *Optom Vis Sci.* 2003;80:832–838.
9. Rouse M, Borsting E, Mitchell G, et al. Validity and reliability of the revised Convergence Insufficiency Symptom Survey in adults. *Ophthalmic Physiol Opt.* 2004;24:384–390.
10. Adler P. Efficacy of treatment for convergence insufficiency using vision therapy. *Ophthalmic Physiol Opt.* 2002;22:565–571.
11. Gallaway M, Scheiman M, Malhotra K. The effectiveness of pencil pushups treatment for

Validation of the SQVD

TVST | February 2022 | Vol. 11 | No. 2 | Article 7 | 10

- convergence insufficiency: a pilot study. *Optom Vis Sci.* 2002;79:265–267.
12. Borsting E, Rouse M, Deland P, et al. Association of symptoms and convergence and accommodative insufficiency in school-age children. *Optometry.* 2003;74:25–34.
 13. Sterner B, Gellerstedt M, Sjöström A. Accommodation and the relationship to subjective symptoms with near work for young school children. *Ophthalmic Physiol Opt.* 2006;26:148–155.
 14. Vaughn W, Maples WC, Hoenes R. The association between vision quality of life and academics as measured by the College of Optometrists in Vision Development Quality of Life questionnaire. *Optometry.* 2006;77:116–123.
 15. Rouse M, Borsting E, Mitchell GL, et al. Academic behaviors in children with convergence insufficiency with and without parent-reported ADHD. *Optom Vis Sci.* 2009;86:1169–1177.
 16. Marella M, Pesudovs K, Keeffe JE, O'Connor PM, Rees G, Lamoureux EL. The psychometric validity of the NEI VFQ-25 for use in a low-vision population. *Invest Ophthalmol Vis Sci.* 2010;51:2878–2884.
 17. Steinberg EP, Tielsch JM, Schein OD, et al. The VF-14. An index of functional impairment in patients with cataract. *Arch Ophthalmol.* 1994;112:630–638.
 18. Lundstrom M, Pesudovs K. Catquest-9SF patient outcomes questionnaire: nine-item short-form Rasch-scaled revision of the Catquest questionnaire. *J Cataract Refract Surg.* 2009;35:504–513.
 19. Kandel H, Pesudovs K, Ferdi A, et al. Psychometric properties of the Keratoconus Outcomes Research Questionnaire: a Save Sight Keratoconus Registry Study. *Cornea.* 2020;39:303–310.
 20. Feinberg DL, Rosner MS, Rosner AJ. Validation of the Binocular Vision Dysfunction Questionnaire (BVDQ). *Otol Neurotol.* 2021;42:e66–e74.
 21. González-Pérez M, Susi R, Antona B, Barrio A, González E. The Computer-Vision Symptom Scale (CVSS17): development and initial validation. *Invest Ophthalmol Vis Sci.* 2014;55:4504–4511.
 22. Seguí MM, Cabrero-García J, Crespo A, Verdú J, Ronda-Pérez E. A reliable and valid questionnaire was developed to measure computer vision syndrome at the workplace. *J Clin Epidemiol.* 2015;68:662–673.
 23. Arroyo-Del Arroyo C, Fernández I, López-de la Rosa A, Pinto-Fraga J, González-García MJ, López-Miguel A. Design of a questionnaire for detecting contact lens discomfort: the Contact Lens Discomfort Index [published online ahead of print March 18, 2021]. *Clin Exp Optom.* doi:10.1080/08164622.2021.1896945.
 24. McAlinden C, Pesudovs K, Moore JE. The development of an instrument to measure quality of vision: the Quality of Vision (QoV) questionnaire. *Invest Ophthalmol Vis Sci.* 2010;51:5537–5545.
 25. McAlinden C, Lipson M. Orthokeratology and Contact Lens Quality of Life Questionnaire (OCL-QoL). *Eye Contact Lens.* 2018;44:279–285.
 26. Pesudovs K, Garamendi E, Elliott DB. The Quality of Life Impact of Refractive Correction (QIRC) Questionnaire: development and validation. *Optom Vis Sci.* 2004;81:769–777.
 27. Hu J, Wang G, Zhou Z, et al. Evaluation of a novel quality of life scale for schoolchildren with non-strabismic binocular vision anomalies. *Biomed Res Int.* 2020;2020:4723402.
 28. Felius J, Beauchamp GR, Stager DR, Sr, Van De Graaf ES, Simonsz HJ. The Amblyopia and Strabismus Questionnaire: English translation, validation, and subscales. *Am J Ophthalmol.* 2007;143:305–310.
 29. Cantó-Cerdán M, Cacho-Martínez P, García-Muñoz A. Delphi methodology for symptomatology associated with visual dysfunctions. *Sci Rep.* 2020;10:19403.
 30. Perneger TV, Courvoisier DS, Hudelson PM, Gayet-Ageron A. Sample size for pre-tests of questionnaires. *Qual Life Res.* 2015;24:147–151.
 31. Cantó-Cerdán M, Cacho-Martínez P, Lara-Lacárcel F, García-Muñoz A. Rasch analysis for development and reduction of Symptom Questionnaire for Visual Dysfunctions (SQVD). *Sci Rep.* 2021;11:14855.
 32. García-Muñoz A, Carbonell-Bonete S, Cantó-Cerdán M, Cacho-Martínez P. Accommodative and binocular dysfunctions: prevalence in a randomised sample of university students. *Clin Exp Optom.* 2016;99:313–321.
 33. Leung YY, Png ME, Conaghan P, Tennant A. A systematic literature review on the application of Rasch analysis in musculoskeletal disease—a special interest group report of OMERACT 11. *J Rheumatol.* 2014;41:159–164.
 34. Kandel H, Khadka J, Watson SL, Fenwick EK, Pesudovs K. Item banks for measurement of refractive error-specific quality of life. *Ophthalmic Physiol Opt.* 2021;41:591–602.
 35. Pesudovs K, Garamendi E, Keeves JP, Elliott DB. The Activities of Daily Vision Scale for cataract surgery outcomes: re-evaluating validity with Rasch analysis. *Invest Ophthalmol Vis Sci.* 2003;44:2892–2899.

Validation of the SQVD

TVST | February 2022 | Vol. 11 | No. 2 | Article 7 | 11

36. Masters GN. A Rasch model for partial credit scoring. *Psychometrika*. 1982;47:149–174.
37. Linacre JM. Optimizing rating scale category effectiveness. *J Appl Meas*. 2002;3:85–106.
38. McAlinden C, Khadka J, de Freitas Santos Paranhos J, Schor P, Pesudovs K. Psychometric properties of the NEI-RQL-42 questionnaire in keratoconus. *Invest Ophthalmol Vis Sci*. 2012;53:7370–7374.
39. Pesudovs K, Burr JM, Harley C, Elliott DB. The development, assessment, and selection of questionnaires. *Optom Vis Sci*. 2007;84:663–674.
40. Andrich D. An expanded derivation of the threshold structure of the polytomous Rasch model that dispels any “threshold disorder controversy”. *Educ Psychol Meas*. 2013;73:78.
41. Andrich D. The legacies of R. A. Fisher and K. Pearson in the application of the polytomous Rasch model for assessing the empirical ordering of categories. *Educ Psychol Meas*. 2013;73:553.
42. Linacre J. What do Infit and Outfit, mean-square and standardized mean? *Rasch Meas Trans*. 2002;16:2:878.
43. Bond T, Fox CM. *Applying the Rasch model: fundamental measurement in the human sciences*. 3rd ed. New York: Routledge; 2015.
44. Stelmack J, Szlyk JP, Stelmack T, et al. Use of Rasch person-item map in exploratory data analysis: a clinical perspective. *J Rehabil Res Dev*. 2004;41:233–241.
45. Christensen KB, Makransky G, Horton M. Critical values for Yen’s Q3: identification of local dependence in the Rasch model using residual correlations. *Appl Psychol Meas*. 2017;41:178–194.
46. Cacho-Martínez P, García-Muñoz A, Ruiz-Cantero MT. Diagnostic validity of clinical signs associated with a large exophoria at near. *J Ophthalmol*. 2013;2013:549435.
47. Khadka J, McAlinden C, Pesudovs K. Quality assessment of ophthalmic questionnaires: review and recommendations. *Optom Vis Sci*. 2013;90:720–744.
48. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull*. 1979;86:420–428.
49. Walter SD, Eliasziw M, Donner A. Sample size and optimal designs for reliability studies. *Stat Med*. 1998;17:101–110.
50. Marx RG, Menezes A, Horovitz L, Jones EC, Warren RF. A comparison of two time intervals for test–retest reliability of health status instruments. *J Clin Epidemiol*. 2003;56:730–735.
51. Wright BD, Linacre JM. Observations are always ordinal; measurements, however, must be interval. *Arch Phys Med Rehabil*. 1989;70:857–860.
52. González-Pérez M, Pérez-Garmendia C, Barrio AR, García-Montero M, Antona B. Spanish cross-cultural adaptation and Rasch analysis of the Convergence Insufficiency Symptom Survey (CISS). *Transl Vis Sci Technol*. 2020;9:23.
53. Davis AL, Harvey EM, Twelker JD, Miller JM, Leonard-Green T, Campus I. Convergence insufficiency, accommodative insufficiency, visual symptoms, and astigmatism in Tohono O’odham students. *J Ophthalmol*. 2016;2016:6963976.
54. Wajuihian SO. Correlations between clinical measures and symptoms: report 2: accommodative and vergence measures with symptoms. *J Optom*. 2021;14:142–155.
55. Scheiman M, Grady MF, Jenewein E, et al. Frequency of oculomotor disorders in adolescents 11 to 17 years of age with concussion, 4 to 12 weeks post injury. *Vision Res*. 2021;183:73–80.
56. Krarup TG, Nisted I, Christensen U, Kiilgaard JF, la Cour M. The tolerance of anisometropia. *Acta Ophthalmol*. 2020;98:418–426.
57. Juniper EF. Validated questionnaires should not be modified. *Eur Respir J*. 2009;34:1015–1017.
58. Tsang S, Royse CF, Terkawi AS. Guidelines for developing, translating, and validating a questionnaire in perioperative and pain medicine. *Saudi J Anaesth*. 2017;11:S80–S89.
59. Horan LA, Ticho BH, Khammar AJ, Allen MS, Shah BA. Is the Convergence Insufficiency Symptom Survey specific for convergence insufficiency? A prospective, randomized study. *Am Orthopt J*. 2015;65:99–103.
60. Horwood AM, Toor S, Riddell PM. Screening for convergence insufficiency using the CISS is not indicated in young adults. *Br J Ophthalmol*. 2014;98:679–683.
61. Clark TY, Clark RA. Convergence Insufficiency Symptom Survey scores for reading versus other near visual activities in school-age children. *Am J Ophthalmol*. 2015;160:905–912.e2.

SECCIÓN 3:

CONCLUSIONES

Universitat d'Alicante
Universidad de Alicante

CONCLUSIONES

Los resultados obtenidos en esta tesis doctoral han permitido alcanzar los objetivos, general y específicos, propuestos.

Como conclusión global, el cuestionario desarrollado y validado en esta investigación, denominado Symptom Questionnaire for Visual Dysfunctions (SQVD), tiene buenas propiedades psicométricas, presenta buena exactitud diagnóstica, y es un instrumento válido y repetible para detectar la presencia y frecuencia de los síntomas relacionados con cualquier tipo de disfunción visual, ya sea refractiva, acomodativa o binocular.

1. La metodología Delphi utilizada ha permitido identificar los síntomas asociados a cualquier tipo de disfunción visual (refractiva, acomodativa y binocular), además de conocer qué síntomas deberían incluirse en un cuestionario según los profesionales de la visión.
 - a. Muchos de los síntomas asignados con mayor frecuencia a las disfunciones refractivas por parte de los expertos, también se asociaron a las disfunciones acomodativas. Además, los expertos asociaron la mayoría de los síntomas a la visión próxima.
 - b. La visión borrosa fue el síntoma con mayor puntuación para las disfunciones refractivas, dificultad para cambios de enfoque para las anomalías acomodativas y ocluir un ojo para las disfunciones binoculares.
 - c. Los expertos participantes en la metodología Delphi alcanzaron una mayoría de acuerdo en que 15 síntomas deberían aparecer en un cuestionario, 5 síntomas no deberían presentarse, existiendo 24

síntomas para los que los expertos tenían discrepancias en cuanto a su inclusión.

2. Utilizando una población clínica se ha desarrollado un cuestionario de sintomatología asociada a cualquier tipo de anomalía visual.
 - a. Se diseñó un cuestionario piloto de 33 ítems utilizando los resultados de la metodología Delphi y aplicando un pretest cognitivo.
 - b. La utilización del análisis Rasch permitió realizar una reducción psicométrica del cuestionario piloto de 33 ítems, obteniendo la versión final de 14 ítems con tres categorías de respuesta. Todos los ítems se ajustaban adecuadamente al modelo Rasch.
3. La utilización del cuestionario SQVD en una muestra clínica mayor ha permitido valorar sus propiedades psicométricas, su exactitud diagnóstica, validez y repetibilidad.
 - a. El SQVD tiene categorías de respuestas ordenadas; los 14 ítems se ajustaron al modelo, no existiendo DIF significativo para las variables de sexo, presbicia y disfunciones; la fiabilidad de la separación de las personas e ítems fue adecuada, al igual que el targeting, mostrando que el cuestionario es unidimensional y no presenta dependencia local entre los ítems.
 - b. El SQVD presenta una buena exactitud diagnóstica. El punto de corte mayor o igual a 6 indica que el paciente presenta síntomas relacionados con una disfunción visual.
 - c. El SQVD tiene una buena validez, mostrando su capacidad para diferenciar sujetos con y sin disfunciones en función de su sintomatología.
 - d. El SQVD presenta muy buena repetibilidad.

SECCIÓN 4:

LÍNEAS FUTURAS DE INVESTIGACIÓN

Las publicaciones científicas de esta tesis doctoral han permitido alcanzar los objetivos presentados al principio de esta investigación. Si bien es cierto que la validación del cuestionario SQVD ha sido exitosa, una de las líneas de investigación futuras sería el análisis de los resultados en una población clínica mayor a la utilizada en la tercera investigación de esta tesis, para así confirmar la mejora de los valores del targeting y de la fiabilidad de los ítems (item separation index).

Otra de las líneas de investigación debería centrarse en el estudio de la validez convergente y validez divergente del cuestionario. Aunque no existe un cuestionario similar que permita valorar la sintomatología asociada a cualquier tipo de disfunción visual, la validez convergente, validez que mide un constructo similar, se podría estudiar comparando el SQVD con otros cuestionarios relacionados con síntomas visuales. Algunos de estos cuestionarios podrían ser el CISS, relacionado con la insuficiencia de convergencia, el cuestionario de Conlon (sobre disconfort visual), el Binocular Vision Dysfunction Questionnaire (BVDQ), relacionado con desviaciones verticales, o el Contact Lens Discomfort Index (CLDI), cuestionario para recoger la sintomatología asociada al disconfort por el uso de lentes de contacto.

Para evaluar la validez divergente, el SQVD se debería comparar con un cuestionario que mida un constructo diferente, como por ejemplo la calidad de visión (Quality of Vision - QoV), calidad de vida (Quality of Life Impact of Refractive Correction - QIRC), o el síndrome visual informático (Computer-Vision Symptom Scale - CVSS17, Computer Vision Syndrome Questionnaire - CSV-Q).

Asimismo, sería interesante ver su comportamiento en muestras clínicas asociadas a consultas hospitalarias y oftalmológicas, al haberse validado en consultas de establecimientos sanitarios de óptica.

Finalmente, en el ámbito clínico, sería necesario comprobar la eficacia del SQVD en la monitorización del tratamiento de las disfunciones visuales, donde se evalúe la evolución de la sintomatología de los pacientes. El SQVD puede ser útil para valorar de forma objetiva la mejoría de los síntomas de los pacientes.



Universitat d'Alacant
Universidad de Alicante

SECCIÓN 5:

BIBLIOGRAFÍA

Universitat d'Alicante
Universidad de Alicante

BIBLIOGRAFÍA

- 1 Sánchez-Cano A, Navarro-Belsué, R. Estado refractivo del ojo y desarrollo de las ametropías. In: Montes-Micó R, editor. *Optometría: Principios básicos y aplicación clínica*. Barcelona: Elsevier España S.L., 2011:3-19.
- 2 Grosvenor T. *Primary Care Optometry*. Boston: Professional Press Books, 1996.
- 3 Eskridge JB, Amos JF, Barlet JD. *Clinical procedure in optometry*. Philadelphia: Lippincott-Williams y Wilkins, 1991.
- 4 Rosenfield M. Refractive Status of the Eye. In: Benjamin WJ, editor. *Borish's Clinical Refraction*. St. Louis: Butterworth-Heinemann, 2006:3-34.
- 5 Scheiman M, Wick B. *Clinical Management of Binocular Vision. Heterophoric, Accommodative, and Eye Movement Disorders*. Wolters Kluwer, 2019.
- 6 Cooper JS, Burns CR, Cotter SA, Daum KM, Griffin JR, Scheiman MM. *Care of the patient with accommodative and vergence dysfunction*. St. Louis: American Optometric Association, 2006.
- 7 Campbell CE, Benjamin WJ, Howland HC. Objective Refraction: Retinoscopy, Autorefractometry and Photorefractometry. In: Benjamin WJ, editor. *Borish's Clinical Refraction*. St. Louis: Butterworth-Heinemann, 2006:683-764.
- 8 Borish IM, Benjamin WJ. Monocular and Binocular Subjective Refraction. In: Benjamin WJ, editor. *Borish's Clinical Refraction*. St. Louis: Butterworth-Heinemann, 2006:791-872.
- 9 Cacho-Martínez P, García-Muñoz A. Evaluación clínica de la visión binocular. In: Montes-Micó R, editor. *Optometría: Principios básicos y aplicación clínica*. Barcelona: Elsevier, 2011:289-317.
- 10 Griffin JR, Grisham JD, Ciuffreda KJ. *Binocular anomalies. Diagnosis and vision therapy*. Amsterdam: Butterworth-Heinemann, 2002.
- 11 Rosenfield M. Clinical assessment of accommodation. In: Rosenfield M, Logan N, Edwards K, editors. *Optometry: science, techniques and clinical management*. Londres: Butterworth-Heinemann, 2009:229-240.
- 12 Hofstetter HW. Useful age-amplitude formula. *Optom World*. 1950;38:42-45.

- 13 Zellers J, Alpert T, Rouse M. A review of the literature and a normative study of accommodative facility. *J Am Optom Assoc.* 1984;55:31-37.
- 14 Ciuffreda KJ. Accommodation, the pupil, and presbyopia. In: Benjamin WJ, editor. *Borish's Clinical Refraction.* St. Louis: Butterworth-Heinemann, 2006:93-144.
- 15 Rouse MW, London R, Allen DC. An evaluation of the monocular estimate method of dynamic retinoscopy. *Am J Optom Physiol Opt.* 1982;59:234-239.
- 16 Saladin JJ. Phorometry and Stereopsis. In: Benjamin WJ, editor. *Borish's Clinical Refraction.* St. Louis: Butterworth-Heinemann, 2006:899-960.
- 17 Morgan MW. The clinical aspects of accommodation and convergence. *Am J Opt Physiol Opt.* 1944;21:301-313.
- 18 Rainey BB, Schroeder TL, Goss DA, Grosvenor TP. Reliability of and comparisons among three variations of the alternating cover test. *Ophthalmic Physiol Opt.* 1998;18:430-437.
- 19 Calvin H, Rupnow P, Grosvenor T. How good is the estimated cover test at predicting the von Graefe phoria measurement? *Optom Vis Sci.* 1996;73:701-706.
- 20 Evans BJW. Binocular vision assessment. In: Rosenfield M, Logan N, editors. *Optometry: Science, Techniques and Clinical Management.* Edinburgh, 2009:241-256.
- 21 Rainey BB, Schroeder TL, Goss DA, Grosvenor TP. Inter-examiner repeatability of heterophoria tests. *Optom Vis Sci.* 1998;75:719-726.
- 22 Rosenfield M. Tonic vergence and vergence adaptation. *Optom Vis Sci.* 1997;74:303-328.
- 23 Rosenfield M. Prism Adaptation: Relevance in clinical practice. *J Optom Vis Dev.* 1997;28:68-75.
- 24 Barnard NA, Thomson WD. A quantitative analysis of eye movements during the cover test--a preliminary report. *Ophthalmic Physiol Opt.* 1995;15:413-419.
- 25 Anderson HA, Manny RE, Cotter SA, Mitchell GL, Irani JA. Effect of Examiner Experience and Technique on the Alternate Cover Test. *Optom Vis Sci.* 2010.

- 26 Johns HA, Manny RE, Fern K, Hu YS. The intraexaminer and interexaminer repeatability of the alternate cover test using different prism neutralization endpoints. *Optom Vis Sci.* 2004;81:939-946.
- 27 Rainey BB, Goss DA, Kidwell M, Feng B. Reliability of the response AC/A ratio determined using nearpoint autorefraction and simultaneous heterophoria measurement. *Clin Exp Optom.* 1998;81:185-192.
- 28 Hayes GJ, Cohen BE, Rouse MW, De Land PN. Normative values for the nearpoint of convergence of elementary schoolchildren. *Optom Vis Sci.* 1998;75:506-512.
- 29 Scheiman M, Gallaway M, Frantz KA et al. Nearpoint of convergence: test procedure, target selection, and normative data. *Optom Vis Sci.* 2003;80:214-225.
- 30 Marran L, De Land P, Nguyen A. Accommodative insufficiency is the primary source of symptoms in children diagnosed with convergence insufficiency. *Optom Vis Sci.* 2006;83:281-289.
- 31 Gallaway M, Scheiman M, Frantz KA, Peters RJ, Hatch S, Cuff M. The significance of assessing near point of convergence using different stimuli. *Optom Vis Sci.* 1991;68 (Supl):93.
- 32 Rouse M, Hyman L, Hussein M, Solan H. Frequency of convergence insufficiency in optometry clinic settings. *Convergence Insufficiency and Reading Study (CIRS) Group.* *Optom Vis Sci.* 1998;75:88-96.
- 33 Rouse M, Borsting E, Hyman L et al. Frequency of convergence insufficiency among fifth and sixth graders. *The Convergence Insufficiency and Reading Study (CIRS) group.* *Optom Vis Sci.* 1999;76:643-649.
- 34 Borsting E, Rouse M, De Land P. Prospective comparison of convergence insufficiency and normal binocular children on CIRS symptom surveys. *Convergence Insufficiency and Reading Study (CIRS) group.* *Optom Vis Sci.* 1999;76:221-228.
- 35 Lanca CC, Rowe FJ. Measurement of fusional vergence: a systematic review. *Strabismus.* 2019;27:88-113.
- 36 Goss DA. Effect of test sequence on fusional vergence ranges. *N Engl J Optom* 1995;47:39-42.

- 37 Daum K. Vergence facility. In: Eskridge JB, Amos JF, Barlet JD, editors. Clinical procedures in optometry. Philadelphia: Lippincott-Williams & Wilkins, 1991:671-676.
- 38 Gall R, Wick B, Bedell H. Vergence facility and target type. *Optom Vis Sci.* 1998;75:727-730.
- 39 Gall R, Wick B, Bedell H. Vergence facility: establishing clinical utility. *Optom Vis Sci.* 1998;75:731-742.
- 40 Ferrer-Blasco T, Madrid-Costa D. Estereopsis. In: Montes-Micó R, editor. *Optometría: Principios básicos y aplicación clínica.* Barcelona: Elsevier España S.L., 2011:145-156.
- 41 Dwyer P. Clinical criteria for vergence accommodation dysfunction. *Clin Exp Optom.* 1991;74:112-119.
- 42 Cacho-Martínez P, García-Muñoz A, Ruiz-Cantero MT. Is there any evidence for the validity of diagnostic criteria used for accommodative and nonstrabismic binocular dysfunctions? *J Optom.* 2014;7:2-21.
- 43 Wick BC. Horizontal deviation. In: Amos J, editor. *Diagnosis and management in vision care.* Boston: Butterworth-Heinemann, 1987:461-510.
- 44 Lara-Lacárcel F, Fernández-Sánchez V. Estrabismos y ambliopía. In: *Optometría: aspectos avanzados y consideraciones especiales.* Barcelona: Elsevier, 2011:55-82.
- 45 Hussaindeen JR, Rakshit A, Singh NK et al. The minimum test battery to screen for binocular vision anomalies: report 3 of the BAND study. *Clin Exp Optom.* 2018;101:281-287.
- 46 Menjivar AM, Kulp MT, Mitchell GL, Toole AJ, Reuter K. Screening for convergence insufficiency in school-age children. *Clin Exp Optom.* 2018;101:578-584.
- 47 Moon BY, Kim SY, Yu DS. Receiver operating characteristic curve analysis of clinical signs for screening of convergence insufficiency in young adults. *PLoS One.* 2020;15:e0228313.
- 48 Cacho-Martínez P, García-Muñoz A, Ruiz-Cantero MT. Diagnostic validity of clinical signs associated with a large exophoria at near. *J Ophthalmol.* 2013;2013:549435.

- 49 García-Muñoz A, Carbonell-Bonete S, Cantó-Cerdán M, Cacho-Martínez P. Accommodative and binocular dysfunctions: prevalence in a randomised sample of university students. *Clin Exp Optom*. 2016;99:313-321.
- 50 Hashemi H, Fotouhi A, Yekta A, Pakzad R, Ostadimoghaddam H, Khabazkhoob M. Global and regional estimates of prevalence of refractive errors: Systematic review and meta-analysis. *J Curr Ophthalmol*. 2018;30:3-22.
- 51 Cacho-Martínez P, García-Muñoz A, Ruiz-Cantero MT. Do we really know the prevalence of accommodative and nonstrabismic binocular dysfunctions? *J Optom*. 2010;3:185-197.
- 52 Shin HS, Park SC, Park CM. Relationship between accommodative and vergence dysfunctions and academic achievement for primary school children. *Ophthalmic Physiol Opt*. 2009;29:615-624.
- 53 Hoseini-Yazdi SH, Yekta A, Nouri H, Heravian J, Ostadimoghaddam H, Khabazkhoob M. Frequency of convergence and accommodative disorders in a clinical population of Mashhad, Iran. *Strabismus*. 2015;23:22-29.
- 54 Wajuihian SO, Hansraj R. Vergence anomalies in a sample of high school students in South Africa. *J Optom*. 2016;9:246-257.
- 55 Wajuihian SO, Hansraj R. Accommodative Anomalies in a Sample of Black High School Students in South Africa. *Ophthalmic Epidemiol*. 2016;23:316-323.
- 56 Hussaindeen JR, Rakshit A, Singh NK et al. Prevalence of non-strabismic anomalies of binocular vision in Tamil Nadu: report 2 of BAND study. *Clin Exp Optom*. 2017;100:642-648.
- 57 Hussaindeen JR, Shah P, Ramani KK, Ramanujan L. Efficacy of vision therapy in children with learning disability and associated binocular vision anomalies. *J Optom*. 2018;11:40-48.
- 58 Hashemi H, Khabazkhoob M, Nabovati P et al. Accommodative insufficiency in a student population in Iran. *J Optom*. 2019;12:161-167.
- 59 Nunes AF, Monteiro PML, Ferreira FBP, Nunes AS. Convergence insufficiency and accommodative insufficiency in children. *BMC Ophthalmol*. 2019;19:58.
- 60 Ma MM, Yeo ACH, Scheiman M, Chen X. Vergence and Accommodative Dysfunctions in Emmetropic and Myopic Chinese Young Adults. *J Ophthalmol*. 2019;2019:5904903.

- 61 Atiya A, Hussaindeen JR, Kasturirangan S, Ramasubramanian S, Swathi K, Swaminathan M. Frequency of undetected binocular vision anomalies among ophthalmology trainees. *J Optom.* 2020;13:185-190.
- 62 Davis AL, Harvey EM, Twelker JD, Miller JM, Leonard-Green T, Campus I. Convergence Insufficiency, Accommodative Insufficiency, Visual Symptoms, and Astigmatism in Tohono O'odham Students. *J Ophthalmol.* 2016;2016:6963976.
- 63 Hussaindeen JR, Murali A. Accommodative Insufficiency: Prevalence, Impact and Treatment Options. *Clin Optom (Auckl).* 2020;12:135-149.
- 64 Hashemi H, Nabovati P, Khabazkhoob M et al. The prevalence of convergence insufficiency in Iran: a population-based study. *Clin Exp Optom.* 2017;100:704-709.
- 65 Hassan LI, Ibrahim SM, Abdu M, MohamedSharif A. Prevalence of convergence insufficiency among secondary school students in Khartoum, Sudan. *Oman J Ophthalmol.* 2018;11:129-133.
- 66 Atowa UC, Wajuihian SO, Hansraj R. Vergence Profile and Prevalance of Non-Strabismic Vergence Anomalies Among School Children in Abia State, Nigeria. *Ophthalmic Epidemiol.* 2019;26:121-131.
- 67 Ma MM, Long W, She Z et al. Convergence insufficiency in Chinese high school students. *Clin Exp Optom.* 2019;102:166-171.
- 68 Oveneri-Ogbomo GO, Eguegu OP. Vergence findings and horizontal vergence dysfunction among first year university students in Benin City, Nigeria. *J Optom.* 2016;9:258-263.
- 69 Scheiman M, Gallaway M, Coulter R et al. Prevalence of vision and ocular disease conditions in a clinical pediatric population. *J Am Optom Assoc.* 1996;67:193-202.
- 70 Fu Z, Hong H, Su Z, Lou B, Pan CW, Liu H. Global prevalence of amblyopia and disease burden projections through 2040: a systematic review and meta-analysis. *Br J Ophthalmol.* 2020;104:1164-1170.
- 71 Hashemi H, Pakzad R, Heydarian S et al. Global and regional prevalence of strabismus: a comprehensive systematic review and meta-analysis. *Strabismus.* 2019;27:54-65.

- 72 Cacho-Martínez P, García-Muñoz A, Ruiz-Cantero MT. Treatment of accommodative and nonstrabismic binocular dysfunctions: a systematic review. *Optometry*. 2009;80:702-716.
- 73 Brautaset R, Wahlberg M, Abdi S, Pansell T. Accommodation insufficiency in children: are exercises better than reading glasses? *Strabismus*. 2008;16:65-69.
- 74 Convergence Insufficiency Treatment Trial Study G. Randomized clinical trial of treatments for symptomatic convergence insufficiency in children. *Arch Ophthalmol*. 2008;126:1336-1349.
- 75 Teitelbaum B, Pang Y, Krall J. Effectiveness of base in prism for presbyopes with convergence insufficiency. *Optom Vis Sci*. 2009;86:153-156.
- 76 Alvarez TL, Vicci VR, Alkan Y et al. Vision therapy in adults with convergence insufficiency: clinical and functional magnetic resonance imaging measures. *Optom Vis Sci*. 2010;87:E985-1002.
- 77 Wahlberg M, Abdi S, Brautaset R. Treatment of accommodative insufficiency with plus lens reading addition: is +1.00 D better than +2.00 D? *Strabismus*. 2010;18:67-71.
- 78 Doble JE, Feinberg DL, Rosner MS, Rosner AJ. Identification of binocular vision dysfunction (vertical heterophoria) in traumatic brain injury patients and effects of individualized prismatic spectacle lenses in the treatment of postconcussive symptoms: a retrospective analysis. *PM R*. 2010;2:244-253.
- 79 Dusek WA, Pierscionek BK, McClelland JF. An evaluation of clinical treatment of convergence insufficiency for children with reading difficulties. *BMC Ophthalmol*. 2011;11:21.
- 80 Kim KM, Chun BY. Effectiveness of home-based pencil push-ups (HBPP) for patients with symptomatic convergence insufficiency. *Korean J Ophthalmol*. 2011;25:185-188.
- 81 Shin HS, Park SC, Maples WC. Effectiveness of vision therapy for convergence dysfunctions and long-term stability after vision therapy. *Ophthalmic Physiol Opt*. 2011;31:180-189.
- 82 Serna A, Rogers DL, McGregor ML, Golden RP, Bremer DL, Rogers GL. Treatment of symptomatic convergence insufficiency with a home-based computer orthoptic exercise program. *J AAPOS*. 2011;15:140-143.

- 83 Scheiman M, Cotter S, Kulp MT et al. Treatment of accommodative dysfunction in children: results from a randomized clinical trial. *Optom Vis Sci.* 2011;88:1343-1352.
- 84 Pang Y, Teitelbaum B, Krall J. Factors associated with base-in prism treatment outcomes for convergence insufficiency in symptomatic presbyopes. *Clin Exp Optom.* 2012;95:192-197.
- 85 Momeni-Moghaddam H, Kundart J, Azimi A, Hassanyani F. The effectiveness of home-based pencil push-up therapy versus office-based therapy for the treatment of symptomatic convergence insufficiency in young adults. *Middle East Afr J Ophthalmol.* 2015;22:97-102.
- 86 Huston PA, Hoover DL. Treatment of symptomatic convergence insufficiency with home-based computerized vergence system therapy in children. *J AAPOS.* 2015;19:417-421.
- 87 Pediatric Eye Disease Investigator G. Home-Based Therapy for Symptomatic Convergence Insufficiency in Children: A Randomized Clinical Trial. *Optom Vis Sci.* 2016;93:1457-1465.
- 88 Rosner MS, Feinberg DL, Doble JE, Rosner AJ. Treatment of vertical heterophoria ameliorates persistent post-concussive symptoms: A retrospective analysis utilizing a multi-faceted assessment battery. *Brain Inj.* 2016;30:311-317.
- 89 Jang JU, Jang JY, Tai-Hyung K, Moon HW. Effectiveness of Vision Therapy in School Children with Symptomatic Convergence Insufficiency. *J Ophthalmic Vis Res.* 2017;12:187-192.
- 90 Nehad T, Salem T, Elmohamady MN. Combined Office-based Vergence Therapy and Home Therapy System for Convergence Insufficiency in Egyptian Children. *Open Ophthalmol J.* 2018;12:12-18.
- 91 Aletaha M, Daneshvar F, Mosallaei M, Bagheri A, Khalili MR. Comparison of Three Vision Therapy Approaches for Convergence Insufficiency. *J Ophthalmic Vis Res.* 2018;13:307-314.
- 92 Mohan A, Kaur N, Sen P, Jain E, Gajraj M. Efficacy of Smartphone-based Exercises in Conjunction with Modified Glasses Prescription in the Treatment of Convergence Insufficiency and Fusion Weakness. *J Binocul Vis Ocul Motil.* 2019;69:30-33.

- 93 Group C-AI. Treatment of Symptomatic Convergence Insufficiency in Children Enrolled in the Convergence Insufficiency Treatment Trial-Attention & Reading Trial: A Randomized Clinical Trial. *Optom Vis Sci.* 2019;96:825-835.
- 94 Group C-AI. Effect of Vergence/Accommodative Therapy on Reading in Children with Convergence Insufficiency: A Randomized Clinical Trial. *Optom Vis Sci.* 2019;96:836-849.
- 95 Nabovati P, Kamali M, Mirzajani A, Jafarzadehpur E, Khabazkhoob M. The effect of base-in prism on vision-related symptoms and clinical characteristics of young adults with convergence insufficiency; a placebo-controlled randomised clinical trial. *Ophthalmic Physiol Opt.* 2020;40:8-16.
- 96 Chen AM, Roberts TL, Cotter SA et al. Effectiveness of vergence/accommodative therapy for accommodative dysfunction in children with convergence insufficiency. *Ophthalmic Physiol Opt.* 2021;41:21-32.
- 97 Singh A, Saxena V, Yadav S et al. Comparison of home-based pencil push-up therapy and office-based orthoptic therapy in symptomatic patients of convergence insufficiency: a randomized controlled trial. *Int Ophthalmol.* 2021;41:1327-1336.
- 98 Alvarez TL, Scheiman M, Santos EM et al. The Convergence Insufficiency Neuro-mechanism in Adult Population Study (CINAPS) Randomized Clinical Trial: Design, Methods, and Clinical Data. *Ophthalmic Epidemiol.* 2020;27:52-72.
- 99 Sangoi A, Scheiman M, Yaramothu C, Santos EM, Gohel S, Alvarez TL. Convergence Insufficiency Neuro-Mechanism Adult Population Study: Phoria Adaptation Results. *Invest Ophthalmol Vis Sci.* 2021;62:19.
- 100 Taylor K, Elliott S. Interventions for strabismic amblyopia. *Cochrane Database Syst Rev.* 2011:CD006461.
- 101 Li T, Qureshi R, Taylor K. Conventional occlusion versus pharmacologic penalization for amblyopia. *Cochrane Database Syst Rev.* 2019;8:CD006460.
- 102 Brin TA, Chow A, Carter C, Oremus M, Bobier W, Thompson B. Efficacy of vision-based treatments for children and teens with amblyopia: a systematic review and meta-analysis of randomised controlled trials. *BMJ Open Ophthalmol.* 2021;6:e000657.

- 103 Pediatric Eye Disease Investigator G, Holmes JM, Manny RE et al. A Randomized Trial of Binocular Dig Rush Game Treatment for Amblyopia in Children Aged 7 to 12 Years. *Ophthalmology*. 2019;126:456-466.
- 104 Roda M, Pellegrini M, Di Geronimo N, Vagge A, Fresina M, Schiavi C. Binocular treatment for amblyopia: A meta-analysis of randomized clinical trials. *PLoS One*. 2021;16:e0257999.
- 105 Asper L, Watt K, Khoo S. Optical treatment of amblyopia: a systematic review and meta-analysis. *Clin Exp Optom*. 2018;101:431-442.
- 106 Chen AM, Cotter SA. The Amblyopia Treatment Studies: Implications for Clinical Practice. *Adv Ophthalmol Optom*. 2016;1:287-305.
- 107 Sen S, Singh P, Saxena R. Management of amblyopia in pediatric patients: Current insights. *Eye (Lond)*. 2022;36:44-56.
- 108 Pang Y, Gnanaraj L, Gayleard J, Han G, Hatt SR. Interventions for intermittent exotropia. *Cochrane Database Syst Rev*. 2021;9:CD003737.
- 109 Kushner BJ. The benefits, risks, and efficacy of strabismus surgery in adults. *Optom Vis Sci*. 2014;91:e102-109.
- 110 Jain S. *Simplifying Strabismus. A Practical Approach to Diagnosis and Management*. Cham, Switzerland: Springer, 2019.
- 111 Cacho-Martínez P, Cantó-Cerdán M, Carbonell-Bonete S, García-Muñoz A. Characterization of Visual Symptomatology Associated with Refractive, Accommodative, and Binocular Anomalies. *J Ophthalmol*. 2015;2015.
- 112 García-Muñoz A, Carbonell-Bonete S, Cacho-Martínez P. Symptomatology associated with accommodative and binocular vision anomalies. *J Optom*. 2014;7:178-192.
- 113 Schellini S, Ferraz F, Opromolla P, Oliveira L, Padovani C. Main visual symptoms associated to refractive errors and spectacle need in a Brazilian population. *Int J Ophthalmol*. 2016;9:1657-1662.
- 114 Marella M, Pesudovs K, Keeffe JE, O'Connor PM, Rees G, Lamoureux EL. The psychometric validity of the NEI VFQ-25 for use in a low-vision population. *Invest Ophthalmol Vis Sci*. 2010;51:2878-2884.

- 115 Steinberg EP, Tielsch JM, Schein OD et al. The VF-14. An index of functional impairment in patients with cataract. *Arch Ophthalmol*. 1994;112:630-638.
- 116 Lundstrom M, Pesudovs K. Catquest-9SF patient outcomes questionnaire: nine-item short-form Rasch-scaled revision of the Catquest questionnaire. *J Cataract Refract Surg*. 2009;35:504-513.
- 117 Kandel H, Pesudovs K, Ferdi A et al. Psychometric Properties of the Keratoconus Outcomes Research Questionnaire: A Save Sight Keratoconus Registry Study. *Cornea*. 2020;39:303-310.
- 118 Arroyo-Del Arroyo C, Fernández I, López-de la Rosa A, Pinto-Fraga J, González-García MJ, López-Miguel A. Design of a questionnaire for detecting contact lens discomfort: the Contact Lens Discomfort Index. *Clin Exp Optom*. 2021:1-7.
- 119 González-Pérez M, Susi R, Antona B, Barrio A, González E. The Computer-Vision Symptom Scale (CVSS17): development and initial validation. *Invest Ophthalmol Vis Sci*. 2014;55:4504-4511.
- 120 Seguí MM, Cabrero-García J, Crespo A, Verdú J, Ronda-Pérez E. A reliable and valid questionnaire was developed to measure computer vision syndrome at the workplace. *J Clin Epidemiol*. 2015;68:662-673.
- 121 McAlinden C, Pesudovs K, Moore JE. The development of an instrument to measure quality of vision: the Quality of Vision (QoV) questionnaire. *Invest Ophthalmol Vis Sci*. 2010;51:5537-5545.
- 122 McAlinden C, Lipson M. Orthokeratology and Contact Lens Quality of Life Questionnaire (OCL-QoL). *Eye Contact Lens*. 2018;44:279-285.
- 123 Pesudovs K, Garamendi E, Elliott DB. The Quality of Life Impact of Refractive Correction (QIRC) Questionnaire: development and validation. *Optom Vis Sci*. 2004;81:769-777.
- 124 Hu J, Wang G, Zhou Z et al. Evaluation of a Novel Quality of Life Scale for Schoolchildren with Nonstrabismic Binocular Vision Anomalies. *Biomed Res Int*. 2020;2020:4723402.
- 125 Felius J, Beauchamp GR, Stager DR, Sr., Van De Graaf ES, Simonsz HJ. The Amblyopia and Strabismus Questionnaire: English translation, validation, and subscales. *Am J Ophthalmol*. 2007;143:305-310.

- 126 Leske DA, Hatt SR, Liebermann L, Holmes JM. Evaluation of the Adult Strabismus-20 (AS-20) questionnaire using Rasch analysis. *Invest Ophthalmol Vis Sci.* 2012;53:2630-2639.
- 127 Feinberg DL, Rosner MS, Rosner AJ. Validation of the Binocular Vision Dysfunction Questionnaire (BVDQ). *Otol Neurotol.* 2021;42:e66-e74.
- 128 Russell GE, Wick B. A prospective study of treatment of accommodative insufficiency. *Optom Vis Sci.* 1993;70:131-135.
- 129 Birnbaum MH, Soden R, Cohen AH. Efficacy of vision therapy for convergence insufficiency in an adult male population. *J Am Optom Assoc.* 1999;70:225-232.
- 130 Conlon EG, Lovegrove WJ, Chekaluk E, Pattison PE. Measuring Visual Discomfort. *Vis Cogn.* 1999;6 (6):637-663.
- 131 Borsting E, Rouse M, Mitchell G et al. Validity and reliability of the revised convergence insufficiency symptom survey in children aged 9 to 18 years. *Optom Vis Sci.* 2003;80:832-838.
- 132 Rouse M, Borsting E, Mitchell G et al. Validity and reliability of the revised convergence insufficiency symptom survey in adults. *Ophthalmic Physiol Opt.* 2004;24:384-390.
- 133 Adler P. Efficacy of treatment for convergence insufficiency using vision therapy. *Ophthalmic Physiol Opt.* 2002;22:565-571.
- 134 Gallaway M, Scheiman M, Malhotra K. The effectiveness of pencil pushups treatment for convergence insufficiency: a pilot study. *Optom Vis Sci.* 2002;79:265-267.
- 135 Borsting E, Rouse M, Deland P et al. Association of symptoms and convergence and accommodative insufficiency in school-age children. *Optometry.* 2003;74:25-34.
- 136 Sterner B, Gellerstedt M, Sjöström A. Accommodation and the relationship to subjective symptoms with near work for young school children. *Ophthalmic Physiol Opt.* 2006;26:148-155.
- 137 Vaughn W, Maples WC, Hoenes R. The association between vision quality of life and academics as measured by the College of Optometrists in Vision Development Quality of Life questionnaire. *Optometry.* 2006;77:116-123.

- 138 Rouse M, Borsting E, Mitchell GL et al. Academic behaviors in children with convergence insufficiency with and without parent-reported ADHD. *Optom Vis Sci.* 2009;86:1169-1177.
- 139 Pesudovs K, Burr JM, Harley C, Elliott DB. The development, assessment, and selection of questionnaires. *Optom Vis Sci.* 2007;84:663-674.
- 140 Khadka J, McAlinden C, Pesudovs K. Quality assessment of ophthalmic questionnaires: review and recommendations. *Optom Vis Sci.* 2013;90:720-744.
- 141 Leung YY, Png ME, Conaghan P, Tennant A. A systematic literature review on the application of Rasch analysis in musculoskeletal disease -- a special interest group report of OMERACT 11. *J Rheumatol.* 2014;41:159-164.
- 142 Wong LP. Focus group discussion: a tool for health and medical research. *Singapore Med J.* 2008;49:256-260; quiz 261.
- 143 Van de Ven AH, Delbecq AL. The nominal group as a research instrument for exploratory health studies. *Am J Public Health.* 1972;62:337-342.
- 144 Black N, Murphy M, Lamping D et al. Consensus development methods: a review of best practice in creating clinical guidelines. *J Health Serv Res Policy.* 1999;4:236-248.
- 145 Dalkey NC. An experimental study of group opinion: The Delphi method. *Futures.* 1969;1:408-426.
- 146 Jones J, Hunter D. Consensus methods for medical and health services research. *BMJ.* 1995;311:376-380.
- 147 Fitch K, Bernstein SJ, Aguilar MD et al. The RAND/UCLA Appropriateness Method User's Manual. RAND, 2001.
- 148 Varela-Ruiz M, Diaz-Bravo L, García-Durán R. Descripción y usos del método Delphi en investigaciones del área de la salud. *Inv Ed Med.* 2012;1(2):90-95.
- 149 Meyrick J. The Delphi method and health research. *Health Educ.* 2003;103:7-16.
- 150 Dalkey NC, Helmer O. An experimental application of the delphi method to the user of experts. *Manage Sci.* 1963;9:458-477.
- 151 Rowe J, Wright G, Bolger F. Delphi: a reevaluation of research and theory. *Technol Forecast Soc Chang.* 1991;39:235-251.

- 152 Pill J. The Delphi method: substance, context, a critique and an annotated bibliography. *Socioecon Plan Sci.* 1971;5 (1):57-71.
- 153 Keeney S, Hasson F, McKenna H. Consulting the oracle: ten lessons from using the Delphi technique in nursing research. *J Adv Nurs.* 2006;53:205-212.
- 154 Powell C. The Delphi technique: myths and realities. *J Adv Nurs.* 2003;41:376-382.
- 155 McMillan SS, King M, Tully MP. How to use the nominal group and Delphi techniques. *Int J Clin Pharm.* 2016;38:655-662.
- 156 Baker J, Lovell K, Harris N. How expert are the experts? An exploration of the concept of 'expert' within Delphi panel techniques. *Nurse Res.* 2006;14:59-70.
- 157 Perneger TV, Courvoisier DS, Hudelson PM, Gayet-Ageron A. Sample size for pre-tests of questionnaires. *Qual Life Res.* 2015;24:147-151.
- 158 Cappelleri JC, Jason Lundy J, Hays RD. Overview of classical test theory and item response theory for the quantitative assessment of items in developing patient-reported outcomes measures. *Clin Ther.* 2014;36:648-662.
- 159 Pesudovs K, Garamendi E, Keeves JP, Elliott DB. The Activities of Daily Vision Scale for cataract surgery outcomes: re-evaluating validity with Rasch analysis. *Invest Ophthalmol Vis Sci.* 2003;44:2892-2899.
- 160 Masters GN. A Rasch model for partial credit scoring. *Psychometrika.* 1982;47.
- 161 Linacre JM. Optimizing rating scale category effectiveness. *J Appl Meas.* 2002;3:85-106.
- 162 McAlinden C, Khadka J, de Freitas Santos Paranhos J, Schor P, Pesudovs K. Psychometric properties of the NEI-RQL-42 questionnaire in keratoconus. *Invest Ophthalmol Vis Sci.* 2012;53:7370-7374.
- 163 Andrich D. An Expanded Derivation of the Threshold Structure of the Polytomous Rasch Model That Dispels Any "Threshold Disorder Controversy". *Educ Psychol Meas.* 2013;73:78.
- 164 Andrich D. The Legacies of R. A. Fisher and K. Pearson in the Application of the Polytomous Rasch Model for Assessing the Empirical Ordering of Categories. *Educ Psychol Meas.* 2013;73:553.

- 165 Linacre J. What do Infit and Outfit, Mean-square and Standardized mean? . Rasch Meas Trans. 2002;16:2:878.
- 166 Christensen KB, Makransky G, Horton M. Critical Values for Yen's Q3: Identification of Local Dependence in the Rasch Model Using Residual Correlations. Appl Psychol Meas. 2017;41:178-194.
- 167 Yasin RM, Yunus FAN, Rus RC, Ahmad A, Rahim RB. Validity and Reliability Learning Transfer Item Using Rasch Measurement Model. Soc Behav Sci. 2014;204:212-217.
- 168 Kandel H, Khadka J, Watson SL, Fenwick EK, Pesudovs K. Item banks for measurement of refractive error-specific quality of life. Ophthalmic Physiol Opt. 2021;41:591-602.
- 169 Bond T. Applying the Rasch model: Fundamental measurement in the human sciences. New York and London: Routledge, 2015.
- 170 Stelmack J, Szlyk JP, Stelmack T et al. Use of Rasch person-item map in exploratory data analysis: a clinical perspective. J Rehabil Res Dev. 2004;41:233-241.
- 171 Marx RG, Menezes A, Horovitz L, Jones EC, Warren RF. A comparison of two time intervals for test-retest reliability of health status instruments. J Clin Epidemiol. 2003;56:730-735.
- 172 Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychol Bull. 1979;86:420-428.
- 173 Chase C, Tosha C, Borsting E, Ridder WH, 3rd. Visual discomfort and objective measures of static accommodation. Optom Vis Sci. 2009;86:883-889.
- 174 Borsting E, Chase C, Tosha C, Ridder WH, 3rd. Longitudinal study of visual discomfort symptoms in college students. Optom Vis Sci. 2008;85:992-998.
- 175 Borsting E, Chase CH, Ridder WH, 3rd. Measuring visual discomfort in college students. Optom Vis Sci. 2007;84:745-751.
- 176 Cooper J, Feldman J. Reduction of symptoms in binocular anomalies using computerized home therapy-HTS. Optometry. 2009;80:481-486.

- 177 Wajuihian SO. Correlations between clinical measures and symptoms: Report 1: Stereoacuity with accommodative, vergence measures, and symptoms. *J Optom.* 2020;13:171-184.
- 178 Wajuihian SO. Correlations between clinical measures and symptoms: Report 2: Accommodative and vergence measures with symptoms. *J Optom.* 2021;14:142-155.
- 179 Scheiman M, Grady MF, Jenewein E et al. Frequency of oculomotor disorders in adolescents 11 to 17 years of age with concussion, 4 to 12 weeks post injury. *Vision Res.* 2021;183:73-80.
- 180 Krarup TG, Nisted I, Christensen U, Kiilgaard JF, la Cour M. The tolerance of anisometropia. *Acta Ophthalmol.* 2020;98:418-426.
- 181 Juniper EF. Validated questionnaires should not be modified. *Eur Respir J.* 2009;34:1015-1017.
- 182 Tsang S, Royse CF, Terkawi AS. Guidelines for developing, translating, and validating a questionnaire in perioperative and pain medicine. *Saudi J Anaesth.* 2017;11:S80-S89.
- 183 Horan LA, Ticho BH, Khammar AJ, Allen MS, Shah BA. Is The Convergence Insufficiency Symptom Survey Specific for Convergence Insufficiency? A Prospective, Randomized Study. *Am Orthopt J.* 2015;65:99-103.
- 184 Horwood AM, Toor S, Riddell PM. Screening for convergence insufficiency using the CISS is not indicated in young adults. *Br J Ophthalmol.* 2014;98:679-683.
- 185 Clark TY, Clark RA. Convergence Insufficiency Symptom Survey Scores for Reading Versus Other Near Visual Activities in School-Age Children. *Am J Ophthalmol.* 2015;160:905-912 e902.
- 186 González-Pérez M, Pérez-Garmendia C, Barrio AR, García-Montero M, Antona B. Spanish Cross-Cultural Adaptation and Rasch Analysis of the Convergence Insufficiency Symptom Survey (CISS). *Transl Vis Sci Technol.* 2020;9:23.
- 187 Dwyer P, Wick B. The influence of refractive correction upon disorders of vergence and accommodation. *Optom Vis Sci.* 1995;72:224-232.
- 188 Chrousos GA, O'Neill JF, Lueth BD, Parks MM. Accommodation deficiency in healthy young individuals. *J Pediatr Ophthalmol Strabismus.* 1988;25:176-179.

- 189 Rutstein R, Daum K, Amos J. Accommodative spasm: a study of 17 cases. *J Am Optom Assoc.* 1988;59:527-538.
- 190 Wright BD, Linacre JM. Observations are always ordinal; measurements, however, must be interval. *Arch Phys Med Rehabil.* 1989;70:857-860.
- 191 Terwee CB, Bot SD, de Boer MR et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol.* 2007;60:34-42.
- 192 Cantó-Cerdán M, Cacho-Martínez P, García-Muñoz A. Delphi methodology for symptomatology associated with visual dysfunctions. *Sci Rep.* 2020;10:19403.
- 193 Walter SD, Eliasziw M, Donner A. Sample size and optimal designs for reliability studies. *Stat Med.* 1998;17:101-110.
- 194 Cantó-Cerdán M, Cacho-Martínez P, Lara-Lacárcel F, García-Muñoz A. Rasch analysis for development and reduction of Symptom Questionnaire for Visual Dysfunctions (SQVD). *Sci Rep.* 2021;11:14855.





SECCIÓN 6:

ANEXOS

Universitat d'Alacant
Universidad de Alicante

Anexo 1. Documento de aprobación de la investigación por el Comité de ética de la Universidad de Alicante.

	<p>Universitat d'Alacant Universidad de Alicante</p>
<p>Vicerectorat d'Investigació i Transferència de Coneixement Vicerrectorado de Investigación y Transferencia de Conocimiento</p>	
<p>AMPARO NAVARRO FAURE, Presidenta del Comitè de Ètica y Vicerectora de Investigación y Transferencia de Conocimiento de la Universidad de Alicante,</p>	<p>AMPARO NAVARRO FAURE, Presidenta del Comitè d'Ètica i Vicerectora d'Investigació i Transferència de Coneixement de la Universitat d'Alacant,</p>
<p>INFORMA:</p>	<p>INFORMA:</p>
<p>Que el Comitè de Ètica de la Universidad de Alicante, reunido en fecha 2 de febrero de 2017, una vez estudiada la documentación presentada del proyecto de Tesis Doctoral "Elaboración y validación de un cuestionario sobre sintomatología asociada a las disfunciones visuales" Exp. UA-2017-1-26 del doctorando Mario Cantó Cerdán, cuyo director y directora de tesis son Ángel García Muñoz y Mª Pilar Cacho Martínez dio su visto bueno para la realización del citado proyecto.</p>	<p>Que el Comitè d'Ètica de la Universitat d'Alacant, reunit amb data 2 de febrer de 2017, una vegada estudiada la documentació presentada del projecte de Tesi Doctoral "Elaboración y validación de un cuestionario sobre sintomatología asociada a las disfunciones visuales" Exp. UA-2017-1-26, del doctorand Mario Cantó Cerdán, el director i la directora de la qual són Ángel García Muñoz y Mª Pilar Cacho Martínez, va donar el seu vistiplau per a la realització de l'esmentat projecte.</p>
<p>Y para que conste a los efectos oportunos se firma el presente en Alicante, a dos de febrero de dos mil diecisiete.</p>	<p>I perquè conste als efectes oportuns es signa el present a Alacant, a dos de febrer de dos mil desset.</p>
<p> Amparo Navarro Faure</p>	<p></p>
<p>Tel. 96 590 3476 - Fax 96 590 9875 Campus de Sant Vicent del Raspeig Ap. 99 E-03080 Alacant e-mail: vr.investi@ua.es web: http://web.ua.es/es/vr-investi</p>	

Anexo 2. Expertos participantes en la metodología Delphi.

Luis Bielsa i Elíes (Cataluña)

Natalia Cano González (País Vasco)

Enrique González Díaz-Obregón (Comunidad de Madrid)

Vicente A. Fernández Sánchez (Región de Murcia)

Francisco Lara Lacárcel (Región de Murcia)

Margarita León López (Comunidad Valenciana)

Rafael Pérez Cambrodí (Comunidad Valenciana)

Mar Pérez-Crespo Gómez (Comunidad de Madrid)

José Miguel Pelegrín Sánchez (Región de Murcia)

David P. Piñero Llorens (Comunidad Valenciana)

Germán Rodríguez Hernández (La Rioja)

Vicente Rodríguez Pascual (Comunidad Valenciana)

Francisco Valderas Jiménez (Comunidad Valenciana)

Teyma Valero Pérez (Comunidad Valenciana)

Anexo 3. Cuestionario Symptom Questionnaire for Visual Dysfunctions (SQVD).

		NO		Ocasionalmente / A menudo	Casi siempre
1	¿Tienes dolor de cabeza al realizar alguna actividad visual?	<input type="checkbox"/>	sí →	<input type="checkbox"/>	<input type="checkbox"/>
2	¿Tienes visión borrosa o desenfocada al realizar una actividad visual en cerca?	<input type="checkbox"/>	sí →	<input type="checkbox"/>	<input type="checkbox"/>
3	¿Se te ponen los ojos rojos o irritados cuando haces una actividad visual?	<input type="checkbox"/>	sí →	<input type="checkbox"/>	<input type="checkbox"/>
4	¿Tienes que inclinar la cabeza o el libro, papel, móvil, tablet, etc., cuando realizas alguna actividad visual?	<input type="checkbox"/>	sí →	<input type="checkbox"/>	<input type="checkbox"/>
5	¿Tienes que releer el texto al realizar alguna actividad de lectura?	<input type="checkbox"/>	sí →	<input type="checkbox"/>	<input type="checkbox"/>
6	¿Te cuesta calcular las distancias cuando fijas la vista en algo?	<input type="checkbox"/>	sí →	<input type="checkbox"/>	<input type="checkbox"/>
7	¿Te lloran los ojos cuando haces una actividad visual?	<input type="checkbox"/>	sí →	<input type="checkbox"/>	<input type="checkbox"/>
8	¿Tienes dificultad para enfocar los objetos cuando cambias de cerca a lejos y/o viceversa?	<input type="checkbox"/>	sí →	<input type="checkbox"/>	<input type="checkbox"/>
9	¿Tienes dolor de cuello o espalda cuando realizas alguna actividad visual?	<input type="checkbox"/>	sí →	<input type="checkbox"/>	<input type="checkbox"/>
10	¿Tienes los ojos secos o con sensación de arenilla cuando haces una actividad visual?	<input type="checkbox"/>	sí →	<input type="checkbox"/>	<input type="checkbox"/>
11	¿Tienes que entornar los ojos para ver mejor?	<input type="checkbox"/>	sí →	<input type="checkbox"/>	<input type="checkbox"/>
12	¿Te entra sueño cuando realizas alguna actividad visual?	<input type="checkbox"/>	sí →	<input type="checkbox"/>	<input type="checkbox"/>
13	¿Tienes dificultad para fijar la mirada al realizar alguna actividad visual?	<input type="checkbox"/>	sí →	<input type="checkbox"/>	<input type="checkbox"/>
14	¿Sientes deslumbramiento o te molesta la luz al realizar alguna actividad visual?	<input type="checkbox"/>	sí →	<input type="checkbox"/>	<input type="checkbox"/>